

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Engenharia de Computação e Automação Industrial

**METODOLOGIA PARA ANÁLISE DE CENAS E
GERAÇÃO DE CENAS PROTOTÍPICAS UTILIZANDO A
TEORIA COMPUTACIONAL DAS PERCEPÇÕES.**

EDNILSON CESAR RODELLA

ORIENTADOR: PROF. DR. ARMANDO FREITAS DA ROCHA

CO-ORIENTADOR: PROF. DR. FERNANDO ANTONIO CAMPOS GOMIDE

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, como parte dos requisitos exigidos para obtenção do título de Mestre em Engenharia Elétrica.

Área de Concentração: Engenharia de Computação.

Banca Examinadora:

Prof. Dr. Eduardo Massad (DIM/FMUSP/USP)

Prof. Dr. Fernando José Von Zuben (DCA/FEEC/UNICAMP)

Prof. Dr. Ricardo Ribeiro Gudwin (DCA/FEEC/UNICAMP)

Campinas – São Paulo – Brasil
2002

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

R614m Rodella, Ednilson Cesar
Metodologia para análise de cenas e geração de cenas
prototípicas utilizando a teoria computacional das
percepções / Ednilson Cesar Rodella.--Campinas, SP:
[s.n.], 2002.

Orientadores: Armando Freitas da Rocha e Fernando
Antonio Campos Gomide.

Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Lógica difusa. 2. Linguagens formais. 3. Visão
por computador. 4. Algoritmos. I. Rocha, Armando
Freitas. II. Gomide, Fernando Antonio Campos. III.
Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. IV. Título.

Resumo

O presente trabalho, visa desenvolver uma metodologia para análise de um conjunto \mathbf{C} , de cenas composta por elementos pertencentes a um conjunto definido de objetos $\{ \mathbf{O}_i \}_{i=1 \text{ to } n}$. O propósito desta análise é (aprender a) identificar as associações mais freqüentes entre estes objetos, que são usados para definir cenas comuns a partir de \mathbf{C} . Essas cenas comuns são chamadas de cenas prototípicas \bar{c} do conjunto de cenas \mathbf{C} . A análise é embasada na Teoria Computacional das Percepções proposta por Zadeh (Zadeh, L.A. "A New Direction in AI – Toward a Computational Theory of Perceptions". In: *AI Magazine*. 2001, 73-84.) e assume que a informação visual no cérebro é apoiada por cadeias de caracteres definidas sobre uma Gramática Fuzzy (Rocha, A.F. "The brain as a symbol processing machine". In: *Progress in Neurobiology*, 1997;53:121-198). Cada \bar{c} é obtida pelas cadeias de caracteres mais freqüentes, definidas por esta Gramática Fuzzy.

Abstract

The present work devotes its attention to the analysis of a set of scenes \mathbf{C} composed by elements belonging to a set of defined objects $\{ \mathbf{O}_i \}_{i=1 \text{ to } n}$. The purpose of this analysis is to (learn to) identify the most frequent associations among these objects, which are used to define the set of common scenes of \mathbf{C} . These common scenes are called prototypical \bar{c} scenes of \mathbf{C} . The analysis is supported by the Computational Theory of Perceptions proposed by Zadeh (Zadeh, L.A. "A New Direction in AI – Toward a Computational Theory of Perceptions". In: *AI Magazine*. 2001, 73-84.) and assumes that visual information in the brain is carried by strings defined over a Fuzzy Grammar (Rocha, A.F. "The brain as a symbol processing machine". In: *Progress in Neurobiology*, 1997;53:121-198). Each \bar{c} is obtained a frequent string in the language defined by this fuzzy grammar.

Agradecimentos

Agradeço em primeiro lugar a Deus, por permitir a realização deste grande sonho e por colocar tantos amigos em meu caminho, que muito me ajudaram na realização deste trabalho.

A meus pais, Olindo e Inês, que através de seu carinho e amor sempre me incentivaram, me orientaram e não mediram esforços para que fosse possível a realização deste sonho. A meu irmão Arthur, que muito me ajudou com sua paciência e simpatia e aos meus familiares que incondicionalmente estiveram comigo em todos os momentos da minha vida.

À minha namorada Luciana Moreno, que foi uma companheira em todos os momentos e que abdicou de inúmeros finais de semana e feriados para me acompanhar, apoiar e incentivar nos estudos e na confecção deste trabalho.

Ao Prof. Dr. Armando Freitas da Rocha, por quem tive o privilégio de ser orientado desde minha iniciação científica e que nunca poupou esforços para me ajudar na realização deste trabalho. Agradeço também, além da orientação, o apoio, companheirismo, experiência e coragem que me passou nesses quatro anos de convivência.

Ao Prof. Dr. Fernando Gomide pelas importantes sugestões e comentários sobre o trabalho.

Aos Professores Doutores Ricardo Ribeiro Gudwin e Fernando Von Zuben, pela dedicação que sempre me receberam e pela indispensável ajuda na correção de alguns problemas desta tese.

A todo o Departamento de Engenharia de Computação e Automação Industrial (DCA) da Faculdade de Engenharia Elétrica e de Computação (FEEC), pela gentileza com que sempre me atenderam e também a todos os professores da FEEC, pelo convívio enriquecedor nos aspectos profissionais e pessoais.

Aos grandes amigos da Eina (Estudos em Inteligência Natural e Artificial), Ana Paula, Andréa, Cecília, Flávia, Luciana, Márcia, Marly, Samanta, que sempre me ajudaram com suas sugestões e incentivos para a realização deste trabalho. Em especial aos desenhistas Cássia e Leonardo, pelas ilustrações utilizadas neste projeto.

Ao amigo de longa data, que considero como um irmão, Fábio Lucchini, cuja contribuição e apoio foram fundamentais para a realização deste trabalho.

Ao amigo e companheiro Mateus Guilherme Fuini, que sempre esteve a meu lado me ajudando e apoiando.

À Alessandra Suhr, pela correção e revisão da dissertação e pelo grande apoio prestado em todas as etapas deste projeto.

Aos doutores Rubens Dutra Alves e Laurinda Matsubashi, que foram grandes incentivadores de meus estudos e que me concederam valiosa lição de vida, além de ajudar a superar inúmeros desafios através de seus sábios e oportunos conselhos.

Aos amigos Maurício, Cristiane, Fábio, Alessandra, Moska, Leslie, Tchê e Ferpa pelos bons momentos de lazer que passamos juntos nos últimos anos e pela torcida e incentivo que sempre depositaram em mim.

Finalmente, meus sinceros agradecimentos à FAPESP, pelo financiamento deste projeto, protocolo 99/105-8.

A todos, os meus mais sinceros e profundos agradecimentos,
MUITÍSSIMO OBRIGADO!

Índice

1	Introdução.....	1
1.1	O sistema visual humano.....	1
1.2	Formalizando o sistema visual humano	3
1.3	Teoria Computacional das Percepções.....	5
1.4	Organização do trabalho.....	7
2	Objetivos do Trabalho	9
2.1	Resumo do Capítulo.....	10
3	Considerações Iniciais.....	11
3.1	Introdução aos Conjuntos e Lógica Fuzzy.....	11
3.1.1	Variável Lingüística.....	12
3.1.2	Função de Pertinência.....	13
3.1.3	Defuzzificação das variáveis lingüísticas.....	14
3.2	Linguagens Formais	16
3.2.1	Alfabeto e Linguagens	16
3.2.2	Gramáticas.....	17
3.2.3	Tipos de Gramáticas.....	20
3.2.4	Gramáticas Livres de Contexto (GLC).....	21
3.3	Linguagem Formal Fuzzy	22
3.4	Sistema Visual Humano	24
3.4.1	Reconhecimento de propriedades básicas das imagens	26
3.4.2	Recomposição de uma cena.....	27
3.4.3	Um Sistema Visual Formal.....	28
3.5	Resumo do Capítulo.....	33
4	Estudo Experimental	35
4.1	Grupo Experimental.....	35
4.2	Ferramentas computacionais.....	40
4.2.1	Editor de Cenários (EdCen)	41
4.2.2	Visualizador e Manipulador de Cenas (VMC).....	44
4.3	Resumo do Capítulo.....	45
5	Geração de cenas prototípicas médias	47
5.1	Algoritmo de promediação convencional para geração da cena média \bar{c}_m	48
5.2	Algoritmo de Promediação Fuzzy para geração da cena prototípica média \bar{c} , utilizando a proposta da TCP.....	50
5.2.1	Obtenção da referência Z	51
5.2.2	Distância de um objeto O_i em relação a uma referência Z	53

5.2.3	Direção de um objeto O_i em relação a uma referência Z	57
5.2.4	Geração da cena prototípica média \bar{c}	61
5.3	Critérios de aceitação de \bar{c}	67
5.4	Resumo do Capítulo.....	68
6	Análise das Descrições I_k das Cenas c_k.....	69
6.1	Resumo do Capítulo.....	73
7	Resultados.....	75
7.1	Abertura	76
7.1.1	Cenas médias \bar{c}_m geradas através da Promediação Convencional.....	76
7.1.2	Cenas prototípicas médias \bar{c} geradas através de Promediação Fuzzy.....	77
7.2	Cena Mais Engraçada	78
7.2.1	Cenas médias \bar{c}_m geradas através da Promediação Convencional.....	79
7.2.2	Cenas prototípicas médias \bar{c} geradas através de Promediação Fuzzy	81
7.2.3	Análise das Descrições I_k das Cenas c_k	83
7.3	Fábrica de Brinquedos.....	85
7.3.1	Cenas médias \bar{c}_m geradas através da Promediação Convencional.....	85
7.3.2	Cenas prototípicas médias \bar{c} geradas através de Promediação Fuzzy.....	87
7.3.3	Análise das Descrições I_k das Cenas c_k	89
7.4	Resumo do Capítulo.....	90
8	Conclusões.....	91
8.1	Sugestões de Trabalhos Futuros	97
9	Referências.....	99
10	Índice Remissivo de Referências.....	105

Índice de Figuras

Figura 1.1 - Localização dos neurônios das vias “o que”, “onde”, responsáveis pela identificação de movimentos e cores.....	1
Figura 3.1 - Exemplo de variável lingüística.	13
Figura 3.2 - Defuzzyficação dos termos lingüísticos de distância.	16
Figura 3.3 - Relação entre os tipos de gramáticas.	21
Figura 3.4 - Etapas da obtenção de uma imagem pelo sistema visual humano.	25
Figura 3.5 - A análise cortical de linhas e ângulo.	26
Figura 3.6 - A análise cortical de cores.	27
Figura 3.7 - A análise cortical do movimento e localização espacial.	27
Figura 3.8 - A especialização hemisférica.	28
Figura 3.9 - Organização computacional dos filtros.	28
Figura 3.10 - Exemplo de um sistema de inteligência artificial, composto por objetos de reconhecimento de elementos, analisando suas características e complexidades. Estes agentes são desenvolvidos para descrever o comportamento dos neurônios do sistema visual natural e segue a mesma hierarquia organizacional das áreas do sistema visual natural.	31
Figura 4.1 - Elementos Pictóricos e Figura de Fundo da atividade “Abertura”.....	38
Figura 4.2 - Elementos Pictóricos e Figura de Fundo da atividade “Cena Mais Engraçada”.....	39
Figura 4.3 - Elementos Pictóricos e figura de fundo da atividade Fábrica de Brinquedos.....	40
Figura 4.4 - Exemplo de montagem de cena no EdCen.	42
Figura 4.5 - Descritor de Cenário.....	43
Figura 4.6 - Visualizador e Manipulador de Cenas.	44
Figura 5.1 - Seleccionador de conjunto de cenas e configurador de limiares de corte.	48
Figura 5.2 - Fluxograma do algoritmo da geração da cena média \bar{c}_m	49
Figura 5.3 - Localização Z	51
Figura 5.4 - Tela de seleção do elemento pictórico O_i que será tomado como referência Z para a geração da cena prototípica média \bar{c}	52
Figura 5.5- Funções de pertinência para a análise da distância entre um objeto O_i e a referência Z	53
Figura 5.6 - Exemplo da classificação de elementos pictóricos em relação à distância.....	54
Figura 5.7 - Exemplo de um histograma de distância de um objeto O_i	55
Figura 5.8 - Gráfico dos histogramas de distâncias.	56
Figura 5.9 - Formulário de montagem das retas fuzzy.....	56
Figura 5.10 - Editando Retas Fuzzy.....	57
Figura 5.11 - Representação gráfica das direções do dicionário D_4	58
Figura 5.12 - Exemplo da classificação da direção de um elemento pictórico O_i em relação a referência Z	59

Figura 5.13 - Exemplo de um histograma de direção de um objeto O_i	60
Figura 5.14 - Gráfico dos histogramas de direções.	61
Figura 5.15 - Exemplos de histogramas de direção e distância de objetos O_i	62
Figura 5.16 - Defuzzyficação das variáveis lingüísticas de distância.	64
Figura 5.17 - Fluxograma do algoritmo de promediação fuzzy.....	66
Figura 5.18 - Imagens com os mesmos objetos, porém com as ordens invertidas.....	68
Figura 6.1 - Analisador de Textos.	69
Figura 6.2 - Exemplo de Dicionário de Termos gerados para as atividades Cena mais Engraçada e Fábrica de Brinquedos.....	71
Figura 6.3 - O dicionário de sintaxe.	71
Figura 7.1 - Cena Média \bar{c}_m original e corrigida por um ser humano com freqüência de uso de objetos O_i superior a 16 indivíduos.....	76
Figura 7.2 - Tela de seleção do elemento pictórico O_r para ser tomado como referência Z na tarefa “Abertura”.....	77
Figura 7.3 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Abertura”, sendo o limiar de corte acima de 16 objetos.....	78
Figura 7.4 - Cena Média \bar{c}_m original e corrigida por um ser humano com freqüência de uso de objetos O_i superior a 11 indivíduos.....	79
Figura 7.5 - Cena Média \bar{c}_m original e corrigida por um ser humano com freqüência de uso de objetos O_i por no mínimo cinco e máximo de 10 indivíduos.	80
Figura 7.6 -Tela de seleção do elemento pictórico O_r para ser tomado como referência Z na tarefa “Cena Mais Engraçada”.	81
Figura 7.7 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Cena Mais Engraçada”, sendo o limiar de corte acima de 11 objetos.	82
Figura 7.8 - Cenas prototípicas médias com limiar de corte acima de 5 objetos e abaixo de 10 objetos.	83
Figura 7.9 - Cena Média original e corrigida por um ser humano Sem Filtros.	86
Figura 7.10 - Cena Média \bar{c}_m original e corrigida por um ser humano com freqüência de uso de objetos O_i superior a 11 indivíduos.....	86
Figura 7.11 - Tela de seleção do elemento pictórico O_r para ser tomado como referência Z na tarefa “Fábrica de Brinquedos”.	87
Figura 7.12 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Fábrica de Brinquedos” sem aplicação de filtragens.....	88
Figura 7.13 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Fábrica de Brinquedos”, sendo o limiar de corte acima de 11 objetos.....	89
Figura 8.1 - Sistema SENSOR.....	95

Índice de Tabelas

Tabela 4.1 - Texto humorístico de Luis Fernando Veríssimo – O Recital.....	36
Tabela 5.1 - Valor de pertinência de cada termo lingüístico a partir da distância euclidiana $d_{i,j}$ entre um objeto O_i e a referência Z	55
Tabela 7.1 - Descrição da Cena Mais Engraçada.	84
Tabela 7.2 - Descrição da Cena Natalina.....	90

1 Introdução

A evolução modelou o homem como um ser visual falante. Isso significa ser capaz de falar sobre o que se está vendo, ou o que foi visto, e visualizar o que é falado. A linguagem fornece uma fala interior para se pensar sobre o mundo visível, enquanto que a imaginação visual permite a visualização do que as palavras estão dizendo.

1.1 O sistema visual humano

Os circuitos neurais envolvidos na linguagem estão distribuídos por quase todo o cérebro, apesar da especialização do hemisfério esquerdo nas funções específicas da linguagem (Bavelier et al., 1997; Bhatnagar et al., 2000; Cabeza & Nyberg, 1997; Cabeza & Nyberg, 2000; Hutsler & Gazzaniga, 1997; Meyer et al., 2000; St George et al., 1999). O mesmo ocorre com a visão, com a diferença da especialização do hemisfério direito na composição mental das imagens exteriores. Esses dois sistemas, o verbal e o visual, trocam informações em certas áreas do lóbulo parietal e frontal (Elias et al., 1999; Gabrieli et al., 1998; Hopfinger et al., 2000; Marshetz et al., 2000; Smith et al., 1998; St George et al., 1999; Ungerleider et al., 1998). Essa troca de informação permite que a linguagem se torne visual e a visão se torne linguagem.

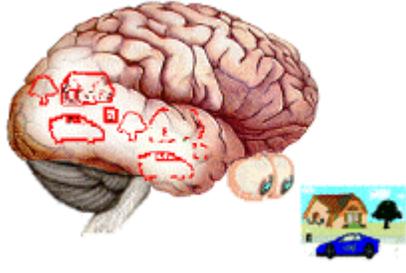
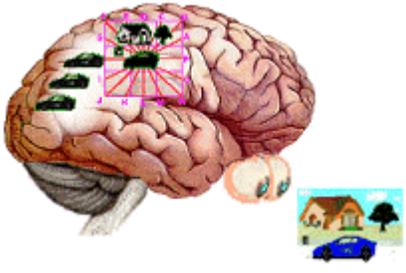
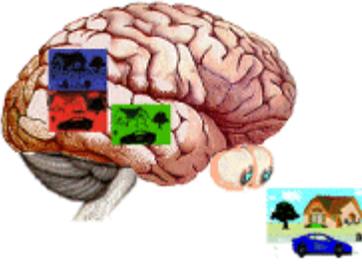
		
<p>Localização dos neurônios das vias “o que”.</p>	<p>Localização dos neurônios das vias “onde” e localização dos neurônios responsáveis pela identificação de movimentos.</p>	<p>Localização dos neurônios responsáveis pela identificação das cores.</p>

Figura 1.1 - Localização dos neurônios das vias “o que”, “onde”, responsáveis pela identificação de movimentos e cores.

O processamento visual depende principalmente da ativação dos neurônios das vias “onde” e “o que” - Figura 1.1 (Ungerleider et al., 1998). As estruturas básicas dos objetos, tais como linhas, orientação dos contornos, comprimento de ondas, etc, ao serem representadas no córtex occipital são processadas por neurônios pertencentes ao processamento visual primário, localizados no córtex

estriado. Neurônios de associação distribuídos ao longo da via “*o que*” combinam essas estruturas básicas para identificar formas mais simples, como triângulos, círculos, cubos, e formas mais complexas, como faces e paisagens (Ishai et al., 2000; Kanwisher & Wojciulik, 2000; Maunsell & Ferrera, 1995; Nakamura et al., 2000; O’Craven & Kanwisher, 2000; Riesenhuber & Poggio, 1999; Young, 1995). A cor é reconhecida principalmente em V1 e V2 e o cálculo de constância de cor é realizado em V4. O terceiro estágio do processamento da cor, baseado na ativação das áreas do córtex frontal e temporal interior, está relacionado com a cor de objetos (Zeki & Marini, 1998). Neurônios associativos, distribuídos ao longo da via “*onde*” no lobo parietal, localizam os elementos identificados pelos neurônios da via “*o que*”, bem como extraem suas relações espaciais, permitindo que as células frontais, hipocampais, etc. reconheçam cenas completas (Johnsrude et al., 1999; Maguire et al., 1998; Nakamura et al., 2000). Portanto, neurônios da via “*o que*” identificam os elementos participantes de uma cena, e os neurônios da via “*onde*”, extraem informações sobre o posicionamento dos elementos numa cena.

A troca de informação verbal e visual nas regiões frontais permite que os dados verbais e visuais sejam integrados, tanto para a identificação total de cenas completas do presente quanto para recriar cenas do passado, além de permitir projetar ações específicas no futuro. A lembrança de eventos pontuais ou prototípicos envolve a participação das áreas frontais verbais - hemisfério esquerdo, e visuais - hemisfério direito (Brewer et al., 1998; Curtis et al., 2000; Ungerleider et al., 1998). Essa troca de mensagem entre os hemisférios frontais permite que a informação verbal seja decodificada em dados visuais e vice-versa, e que uma memória episódica verbal e/ou visual seja acionada em cada hipocampo e suas áreas adjacentes (Brewer et al., 1998; Sawrie et al., 2000).

Uma série de estudos vem propondo que a imaginação mental possa recrutar muitos dos mecanismos envolvidos na percepção visual (Cabeza & Nyberg, 1997; Cabeza & Nyberg, 2000; Kosslyn et al., 1999; Mellet et al., 2000; Perani et al., 1999; Raji, 1999), embora exista a discussão se as áreas de processamento visual primário participam da criação das imagens mentais (Mellet et al., 2000). Palavras concretas parecem recrutar muito mais neurônios das áreas visuais que as palavras abstratas (West & Holcomb, 2000). Em geral, os indivíduos também codificam e produzem mais rápida e completamente palavras concretas que abstratas. Vem sendo proposto que são as palavras concretas, e não as abstratas, as capazes de ativar um sistema “imagístico”, e isso daria uma vantagem no reconhecimento das palavras concretas em relação às abstratas (Holcomb et al., 1999; Holcomb & McPherson, 1994; Paivio, 1996; Paivio, 1991; West & Holcomb, 2000).

Rocha e colaboradores (Rocha et al., 2002) desenvolveram uma técnica de Mapeamento Cognitivo Cerebral (MCC) a partir do registro do eletroencefalograma, e mostraram que a compreensão de um texto é modificada pelo uso ou não de imagens que ilustrem a história contada. Um conto humorístico de L. F. Veríssimo foi gravado por um falante nativo do português e também ilustrado por um desenhista profissional. Apresentou-se o texto falado e ilustrado para um grupo (A) de indivíduos e

somente o texto falado para outro grupo (B) experimental, enquanto registrava-se o eletroencefalograma do voluntário. Depois, o indivíduo era submetido a tarefas para avaliar sua compreensão da história. Os resultados mostraram que na ausência da ilustração os indivíduos imaginam a história contada, utilizando o córtex frontal para recodificar a informação verbal em visual e coordenar áreas parieto-occipital para mentalmente visualizar a cena imaginada. O uso de ilustração modifica o processamento, porque os indivíduos são mais verbais ou visuais no tratamento da informação. Um indivíduo verbal apóia-se mais sobre o processamento lingüístico e utiliza a informação visual para corroborar os resultados desse processamento verbal. O contrário ocorre com os indivíduos mais visuais, que utilizam a informação verbal para complementar o processamento visual.

Esses resultados mostram que os MCCs não só identificam adequadamente as áreas cerebrais mais envolvidas na solução de uma tarefa cognitiva, mas também demonstram o caráter distribuído do processamento cerebral e a dependência da inteligência em relação a essa capacidade de recrutamento neural (Rocha, 1999; Rocha et al., 2002).

1.2 Formalizando o sistema visual humano

Para a obtenção de uma imagem, o sistema visual humano necessita dos seguintes agentes:

- Um conjunto **R** de agentes primários localizados na retina, que extraem informações básicas sobre luminância, cores, movimento e contraste em uma imagem **I**;
- Um conjunto **V** de agentes corticais primários, que utilizam as informações fornecidas por **R** para reconhecer atributos primários da imagem, tais como, linhas, ângulos, descontinuidades de linhas, etc. na imagem **I**,
- Um conjunto **E** de agentes corticais que utilizam as informações fornecidas por **V** para reconhecer objetos na imagem **I**.

Usando as teorias de Sistema Inteligente de Processamento Distribuído (SIPD), e de Linguagem Formal Nebulosa, Rocha e colaboradores (Serapião et al., 1996; Rocha, 1997) propuseram uma formalização do sistema visual humano e geraram um sistema denominado **SENSOR** capaz de simular o processamento visual até o reconhecimento de objetos, inclusive com capacidade de aprender a fazer esse reconhecimento a partir de um conjunto de exemplos.

O sistema **SENSOR** foi utilizado com sucesso por aplicativos para a indústria de petróleo, reconhecimento automatizado de padrões com o propósito de controle (Alegre et al., 1993a,b), entre outros.

Lucchini (Lucchini, 2001) também utilizou o conceito proposto pelo sistema **SENSOR** para o desenvolvimento do sistema **REMMC** (Reconhecimento especial de marcadores para manuseio de computadores), o qual analisa e processa imagens capturadas por uma câmera de vídeo em tempo real para obtenção de características dos elementos presentes na imagem, a fim de localizar um elemento previamente definido, elemento este chamado de marcador, com a finalidade de rastrear seu movimento dentro da imagem, cujo movimento detectado é usado para controlar o ponteiro do mouse de um computador, sem a necessidade de utilizar o periférico mouse. Assim sendo, pessoas portadoras de deficiências motoras, podem utilizar o computador, colocando um adesivo (marcador) na parte do corpo que o indivíduo tenha maior controle motor.

O próximo passo do aprimoramento do **SENSOR** requer o desenvolvimento da formalização de um conjunto **O** de agentes que sejam capazes de analisar as características de cenas na imagem **I**, compostas pelos objetos reconhecidos por agentes $e \in E$, e possam reconhecer cenas prototípicas.

Pode-se propor que, dadas cenas c_k compostas com o auxílio de um editor de cena, que disponibilize um conjunto de elementos **E**, as relações entre esses elementos possam ser obtidas a partir de suas relações de direção e distância. Desta forma, estaríamos simulando os neurônios da via “onde”.

Para diminuir a complexidade deste projeto, o editor de cenas utilizado não usará o sistema **SENSOR** para identificação dos elementos presentes nas cenas. Ao invés disso, o editor de cenas será vetorial, guardando informações sobre os elementos tais como posicionamento e características (cor, forma, etc). Essas informações serão armazenadas numa base de dados específica. Desta forma será possível calcular a distância e angulação entre elementos, já que é possível acessar a posição de cada elemento na tela.

Além da análise das relações espaciais entre objetos, o sistema também será capaz de adquirir conhecimento a partir da análise de um grupo de cenas semelhantes. Este conhecimento será obtido através da montagem de cenas prototípicas, que serão montadas a partir da análise de posicionamento de objetos de um grupo de cenas semelhantes, sendo que cenas semelhantes são cenas geradas sobre um mesmo tema, como por exemplo um tema natalino.

As cenas prototípicas conterão os objetos que se encaixarem nas restrições de frequência definidas, e estes objetos serão colocados nas direções e distâncias de maior frequência. Portanto, podemos dizer que o conhecimento será extraído de acordo com o conjunto de restrições definidas.

1.3 Teoria Computacional das Percepções

Os seres humanos possuem uma notável capacidade de realizar uma grande variedade de tarefas físicas e mentais sem a necessidade de medições ou cálculos precisos. Os valores dos atributos são granulados e cada grânulo reúne um grupo de valores (pontos, objetos) por similaridade, proximidade ou função. Por exemplo, os grânulos de uma distância podem ser marcados como muito perto, perto, médio, longe, muito longe, dentre outros. Quando esses grânulos têm limites imprecisos, são chamados F-Grânulos (Zadeh, 2001).

A Teoria Computacional de Percepções (TCP) proposta por Zadeh (Zadeh, 2001) como um paradigma do raciocínio impreciso, pressupõe que:

1. As percepções humanas são organizadas a partir de F-Grânulos;
2. São muitas vezes descritas em linguagem natural, como por exemplo, o objeto X está longe (perto...) e aproximadamente à direita (esquerda...) do objeto Y.

Na TCP, uma proposição p é vista como uma resposta a uma questão q e o significado de p é representado como uma restrição generalizada R . O cálculo com percepções é proposto como um processamento de uma Linguagem de Restrições Generalizadas (LRG), de modo que a propagação das restrições tem como meta responder uma determinada questão (Zadeh, 2001).

Por exemplo, o processamento da análise de uma cena visual põe questões q do tipo:

“Qual a localização de X em relação a Y?”

que requer como solução dois tipos de proposições:

- p_1 : sobre as relações de distância entre os objetos que compõe a cena e
- p_2 : sobre as direções entre esses mesmos objetos.

Em outras palavras, o reconhecimento visual de uma cena c_k pode ser feito respondendo a uma questão q referente a localização de objetos que compõe c_k :

$$(I) \quad q = (p_1, p_2) = \text{Localização}(O_x, O_y)$$

E a questão q poderia ter duas respostas:

- (II) p_1 : A distância entre (O_x, O_y) é {Muito Perto, Perto, Média, Longe, Muito Longe}
 p_2 : A direção entre (O_x, O_y) é {Norte, Sul, Leste, Oeste, Noroeste...}

De modo que a solução é encontrada com o tratamento das restrições:

- (III) $R_1 = \{\text{Muito Perto, Perto, Média, Longe, Muito Longe}\}$
 $R_2 = \{\text{Norte, Sul, Leste, Oeste, Noroeste...}\}$

A análise completa de c_k , realizada como cálculo da localização do conjunto de objetos O_i em relação a uma referência Z , poderia ser formalizada do seguinte modo:

- (IV) $c_k = [q_1, q_2, \dots, q_n]$, onde:
 $q_i = (p_1, p_2) = \text{Localização}(O_i, Z)$

o que significa que a questão q é repetida para todos os objetos presentes na cena c_k . A referência Z pode ser o centro de um objeto específico O_r ou um ponto qualquer no espaço XY .

Dado um conjunto $C = \{c_1, c_2, \dots, c_m\}$ de cenas semelhantes, é possível calcular uma cena prototípica \bar{c} composta pelos objetos O_i mais freqüentes em C , sendo que consideramos cenas semelhantes, cenas que possuam o mesmo tema, e que os objetos que as compõem, façam parte de um mesmo grupo de elementos. Porém, não é necessário que uma cena contenha todo o grupo de objetos.

Um exemplo de um conjunto C de cenas semelhantes, seria um grupo de cenas sobre o tema “Natal”, e o grupo de objetos que poderiam ser usados nestas cenas seriam por exemplo: “Papai Noel”, “Trenó”, “Cartas”, etc.

Podemos obter a cena prototípica \bar{c} através de:

- (V) $\bar{c} = \text{Prot1}(C)$

onde **Prot1** é um algoritmo capaz de gerar uma cena prototípica \bar{c} , a partir de um conjunto C de cenas semelhantes. A cena prototípica \bar{c} será caracterizada pela disposição dos objetos O_i em relação a Z nas cenas que compõem C , ordenada por sua freqüência em C , isto é, os objetos mais freqüentes terão prioridade no seu posicionamento em \bar{c} . O algoritmo **Prot1** será explorado no capítulo 5 - Geração de cenas prototípicas médias.

A freqüência de posicionamento de objetos pode ser obtida a partir dos histogramas dos possíveis valores assumidos pelas proposições p_1 e p_2 geradas para cada um dos objetos O_i em C .

Esses histogramas informarão os valores assumidos por cada termo lingüístico de distância (Muito Perto, Perto, Média, Longe, Muito Longe) e direção (Norte, Sul, Leste, Oeste, Noroeste,...). Pode-se supor que as proposições que obtiveram maior freqüência em cada histograma correspondem a um consenso sobre a localização dos objetos O_i em C , e por isso devem definir \bar{c} .

Outra proposição básica da TCP é que as percepções são descritas com linguagem natural. Seja I_k uma descrição da cena c_k , pode-se propor que I_k seja um conjunto de sentenças aceitas por uma gramática G . Ou seja:

$$(VI) \quad I_k = \{s_1, s_2, \dots, s_n\}, \text{ onde } s_i \text{ é uma sentença pertencente a uma gramática } G$$

O conceito de gramática será melhor abordado no capítulo 3 - Considerações Iniciais.

Dado um conjunto $L = \{I_1, I_2, \dots, I_m\}$ de descrições, Rocha e colaboradores (Rocha, 1992; Rocha et al., 1992; Guilherme, 1996), propuseram obter uma descrição \bar{I} de consenso, como:

$$(VII) \quad \bar{I} = \text{Prot2}(L)$$

onde **Prot2** é um algoritmo que tenta obter uma descrição \bar{I} de consenso, a partir de um conjunto L de descrições de cenas. A frase \bar{I} será composta pelas sentenças mais freqüentes de L .

1.4 Organização do trabalho

Este trabalho está organizado da seguinte forma:

No capítulo 2 serão apresentados os objetivos do presente trabalho.

No capítulo 3 será apresentado um breve resumo sobre alguns aspectos teóricos necessários para a compreensão deste trabalho. Como por exemplo: Linguagens Formais, Conjuntos e Lógica Fuzzy, Linguagens Formais Nebulosas e o Sistema Visual Humano.

O capítulo 4 tratará dos procedimentos utilizados na pesquisa, tais como o grupo de voluntários selecionados, as atividades que estes voluntários desenvolveram e os softwares criados para que pudesse ser possível a realização deste trabalho.

No capítulo 5, serão mostrados os algoritmos para a geração de Cenas Médias \bar{c}_m e Cenas Prototípicas Médias \bar{c} .

O capítulo 6 mostrará como é feita a análise das descrições das cenas desenvolvidas pelos voluntários.

No capítulo 7, serão apresentados os resultados obtidos com a pesquisa.

O capítulo 8 diz respeito às conclusões obtidas através da pesquisa realizada, e no capítulo 9 encontram-se as referências bibliográficas citadas neste trabalho.

2 Objetivos do Trabalho

O propósito do presente trabalho é utilizar a TCP (Teoria Computacional das Percepções) para estudar:

- 1) A criação de cenas c_k sobre um tema específico ou orientada por uma história;
- 2) As descrições l_k , formada por um conjunto de sentenças dessas cenas c_k ;
- 3) Obtenção de cenas prototípicas médias \bar{c} e as correspondentes descrições \bar{l} de consenso.

Para esse fim, deve-se utilizar um editor de desenhos que permita o uso de elementos pictóricos pré-definidos, que sejam correlacionados com o tema ou história escolhida e a utilização de uma linguagem de programação de objetos, de modo que a cena c_k possa ser construída visualmente e descrita verbalmente. Esses elementos pictóricos, são objetos O_i na forma de imagens que ao serem colocados sobre uma figura de fundo, formam uma cena c_k .

Ou seja, c_k pode ser:

- Caracterizada visualmente como na equação (IV):

$$c_k = [q_1, q_2, \dots, q_n], \text{ onde:}$$
$$q_i = (p_1, p_2) = \text{Localização}(O_i, Z)$$

- Descrita verbalmente como na equação (VI):

$$l_k = \{s_1, s_2, \dots, s_n\}, \text{ onde}$$

s_i é uma sentença pertencente a uma gramática G

A adequação de \bar{c} será validada pela sua comparação com uma cena \bar{c}_m obtida pela promediação de posições dos objetos utilizados em um conjunto C formado por cenas semelhantes c_k e principalmente, por produzir ou não um conjunto visualmente aceitável por parte de qualquer pessoa.

A princípio, uma cena é visualmente aceitável se os objetos que a compõe seguem certas regras gerais de posicionamento. Por exemplo, em uma cena visualmente aceitável, uma casa não pode estar

localizada dentro de um lago, ou no céu. Assim como uma casa não pode estar localizada em cima de outra.

Desta forma, pode ser observado que a figura de fundo de uma cena, possui um importante papel na definição das regras de posicionamento que um indivíduo usa para definir se uma cena é ou não visualmente aceitável.

As regras de posicionamento serão melhor abordadas no item 5.3 – Critérios de aceitação de \bar{c} .

A correlação em \bar{c} e \bar{l} será subjetivamente avaliada pela aceitação ou não de \bar{l} como uma descrição de \bar{c} , aceitável em Português.

A validação da TCP para análise de cenas fornecerá a ferramenta para formalização dos agentes $\mathbf{o} \in \mathbf{O}$ que sejam capazes de aprender a identificar cenas prototípicas e, portanto, permitir expandir o sistema **SENSOR**.

Além disso, essa validação também servirá de evidência da proposição de Lotfi A. Zadeh, de que a TCP seria o mecanismo básico para simulação do raciocínio humano.

2.1 Resumo do Capítulo

Neste capítulo, foram apresentados os objetivos da realização deste trabalho.

No capítulo seguinte, será apresentado sinteticamente, alguns aspectos teóricos necessários para a compreensão deste trabalho.

3 Considerações Iniciais

Neste capítulo, será apresentado um breve resumo sobre alguns aspectos teóricos necessários para a compreensão deste trabalho. Como por exemplo: Conjuntos e Lógica Fuzzy, Linguagens Formais, Linguagens Formais Nebulosas e o Sistema Visual Humano.

3.1 Introdução aos Conjuntos e Lógica Fuzzy

Um dos mais importantes pontos na evolução do conceito moderno sobre incerteza foi a publicação do artigo *Fuzzy Sets*, por Lotfi A. Zadeh em 1965 (Zadeh, 1965). Neste artigo, Zadeh modificou a teoria de conjuntos, na qual muitos fundamentos são baseados na lógica aristotélica, onde uma premissa ou é completamente verdadeira, ou completamente falsa, isto é, um elemento pertence ou não pertence a um determinado conjunto.

As aplicações dos conjuntos fuzzy, alteram a fundamentação da lógica aristotélica por considerar o conceito de verdade parcial. Quando A é um conjunto fuzzy e x é um objeto relevante, a proposição “ x é um membro de A ”, pode ter um certo grau de verdade, não sendo necessariamente verdadeiro ou falso, como seria se estivéssemos utilizando a lógica aristotélica. É muito comum expressarmos graus de pertinência em um conjunto fuzzy como graus de verdade associados a proposições, pertencentes ao intervalo fechado $[0, 1]$, onde os valores extremos significam respectivamente a falsidade ou a veracidade total de uma proposição.

Os conjuntos fuzzy nos proporcionam, além de uma poderosa representação de incerteza de medidas, uma significativa representação de conceitos vagos expressados em linguagem natural.

Serão usadas neste trabalho, as seguintes notações para representação da distância entre um objeto O_i e uma referência Z qualquer: $d_{i,j}$ (em negrito e itálico), para representação de uma distância através de um termo lingüístico, e $\mathbf{d}_{i,j}$ (somente em negrito) para representação da distância numérica (crisp). Ou seja:

(VIII) $d_{i,j}$ = termo lingüístico que representa a distância entre um objeto O_i e uma referência Z .
 $\mathbf{d}_{i,j}$ = distância euclidiana entre um objeto O_i e uma referência Z .

Variável Lingüística

As percepções humanas de medidas são feitas de forma imprecisa e geralmente no formato de linguagem natural (Zadeh, 2001), como por exemplo, a temperatura está quente, o preço é caro, o objeto O_1 esta perto do objeto O_2 .

Uma Variável Lingüística (Zadeh, 1975; Zadeh, 1976) é um tipo de construção capaz de representar conceitos lingüísticos tais como: Muito Perto, Perto, Médio, Longe, Muito Longe, e assim por diante.

Os estados de cada variável lingüística são expressos por um termo lingüístico, definidos através de uma variável base. A variável base é uma variável clássica que representa valores físicos – crisp (Klir & Yuan, 1995), como por exemplo, a distância euclidiana $d_{i,j}$ entre o centro de um objeto O_i e o centro de uma referência Z . Já na variável lingüística, termos lingüísticos representam valores aproximados da variável base.

Cada variável lingüística é caracterizada por um quintuplo (ν, T, X, g, m) , onde ν é o nome da variável, T é o conjunto dos termos lingüísticos de ν que referem à variável base, no qual a amplitude dos valores é coberto pelo conjunto Universo X , g é uma gramática para geração dos termos lingüísticos e m é uma regra de semântica que nomeia cada termo lingüístico $t \in T$, onde $m(t)$ é um conjunto fuzzy em X (Klir & Yuan, 1995; Zimmermann, 1996).

Um exemplo de variável lingüística é apresentado na Figura 3.1. O nome da variável lingüística é *distância*, e ela expressa a distância entre um objeto O_i e uma referência Z através de cinco termos lingüísticos: Muito Perto, Perto, Médio, Longe, Muito Longe.

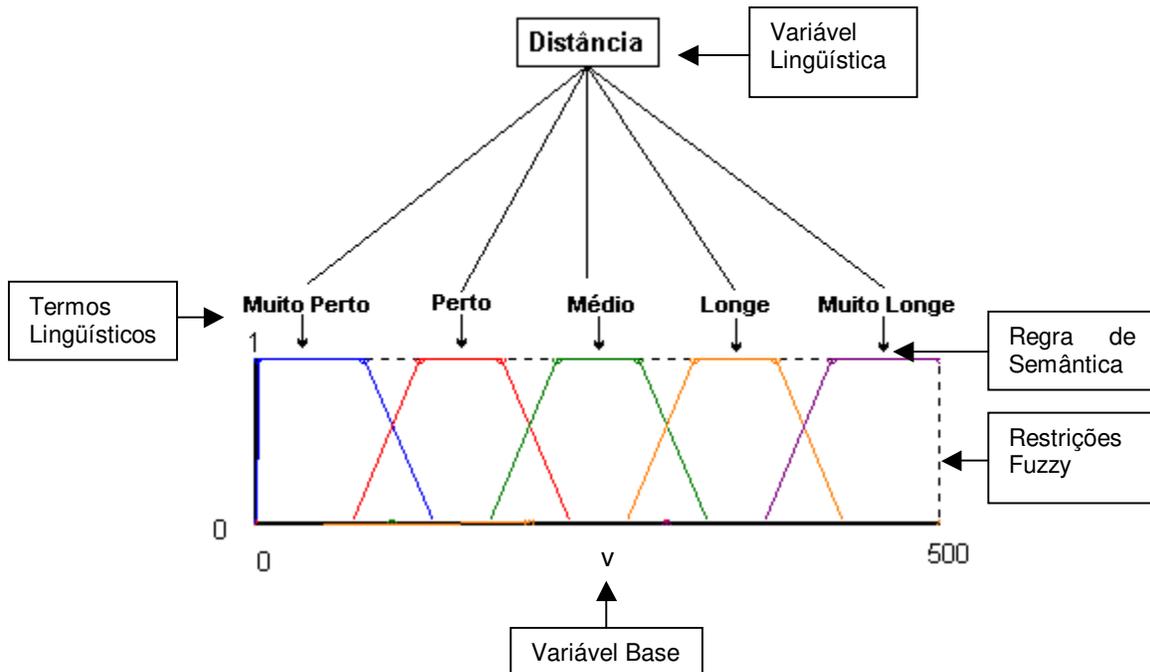


Figura 3.1 - Exemplo de variável lingüística.

A significância das variáveis lingüísticas é a facilidade que elas proporcionam para transições graduais entre estados e, conseqüentemente, a capacidade natural de expressão de medidas incertas (Zadeh, 2001, Klir & Yuan, 1995).

3.1.1 Função de Pertinência

Apesar das variáveis lingüísticas proporcionarem medidas mais vagas e menos específicas do que medidas matemáticas clássicas, são mais usáveis e universais, pois são adaptativas ao contexto. Já medidas “crisp” (clássicas), não são universais.

Por exemplo, R\$ 10,00 pode ser “muito barato” para um banqueiro e “muito caro” para um estudante. Portanto, a premissa “se o objeto custar R\$ 10,00, não compre”, não funcionará igualmente para o banqueiro e para o estudante, mas se a premissa for “se o objeto custar caro, não compre”, funcionará para ambos os casos a despeito das diferenças econômicas, culturais e lingüísticas.

Como se pode observar, portanto, um conjunto fuzzy pode ser definido matematicamente, nomeando para cada possível indivíduo neste universo de discurso, um valor que representa seu grau de pertinência neste conjunto fuzzy. Este grau de pertinência indica o grau com que um indivíduo pertence

ou não a um determinado conjunto fuzzy, podendo ser caracterizado por uma função denominada função de pertinência ou função de membership (Zadeh, 1965; Klir & Yuan, 1995; Zimmermann, 1996).

Como mencionado anteriormente, a amplitude de uma função de pertinência, encontra-se geralmente no intervalo entre 0 a 1. Desta forma cada função de pertinência, mapeia elementos de um dado conjunto X em um número real no intervalo $[0,1]$.

No presente trabalho, utilizaremos a seguinte notação para representar uma função de pertinência:

(IX) $\mu_A : X \rightarrow [0,1]$, onde A representa o nome do conjunto fuzzy.

A Figura 3.1, mostra as funções de pertinência usadas para definir o valor de uma variável fuzzy, a partir de um valor referente à distância entre um objeto O_i e uma referência Z qualquer.

Uma função de pertinência não precisa necessariamente ter a forma trapezoidal mostrada na Figura 3.1. Cada aplicação pode ter uma função de pertinência na forma que melhor se adapte ao contexto trabalhado. As formas trapezoidais e triangulares são as mais usadas.

3.1.2 Defuzzificação das variáveis lingüísticas

O propósito da defuzzificação é converter o conteúdo de uma variável lingüística, que é expressa em termos de um conjunto fuzzy, para um único número real (Klir & Yuan, 1995).

Os métodos de defuzzificação predominantes na literatura são (Klir & Yuan, 1995; Zimmermann, 1996): (1) Método do Centro da Área (Método do Centro de Gravidade); (2) Método do Centro dos Máximos; (3) Método da Média dos Máximos.

Neste trabalho foi usada a seguinte metodologia para o processo de defuzzificação da distância expressa em termos lingüísticos $d_{i,j}$, para a distância euclidiana $d_{i,j}$:

Inicialmente, é feito o seguinte cálculo:

(X) $(q_{i-1} * \mu_{i-1} + q_i * \mu_i + q_{i+1} * \mu_{i+1}) / (\mu_{i-1} + \mu_i + \mu_{i+1})$

Onde:

- μ_i , é o valor da função de pertinência de maior grau.
- q_i , é o ponto médio da região de máximo valor da função de pertinência de maior grau (μ_i).
- i , é o termo lingüístico referente à função de pertinência de maior grau. Os valores que i pode assumir são: “Muito Perto”, “Perto”, “Médio”, “Longe”, “Muito Longe”.

Conseqüentemente, $q_{i-1} * \mu_{i-1}$ diz respeito ao cálculo da função de pertinência referente ao termo lingüístico anterior ao termo lingüístico de maior grau de pertinência, multiplicado pelo ponto médio da região de máximo valor da função de pertinência (μ_{i-1}). Se $i = \text{“Muito Perto”}$, esta etapa é ignorada.

Já, $q_{i+1} * \mu_{i+1}$ diz respeito ao cálculo da função de pertinência referente ao termo lingüístico posterior ao termo lingüístico de maior grau de pertinência, multiplicado pelo ponto médio da região de máximo valor da função de pertinência (μ_{i+1}). Se $i = \text{“Muito Longe”}$, esta etapa é ignorada.

Portanto, supondo que a função de pertinência de maior grau fosse referente ao termo lingüístico *Médio*, a defuzzificação seria feita da seguinte forma:

$$(q_{\text{Perto}} * \mu_{\text{Perto}} + q_{\text{Médio}} * \mu_{\text{Médio}} + q_{\text{Longe}} * \mu_{\text{Longe}}) / (\mu_{\text{Perto}} + \mu_{\text{Médio}} + \mu_{\text{Longe}})$$

Devido ao formato das funções de pertinência usadas neste trabalho, sempre que o grau de pertinência for igual a um, o resultado da defuzzificação será igual a q_i , já que μ_{i-1} e μ_{i+1} terão o valor zero.

A Figura 3.2, mostra as funções de pertinência de cada um dos cinco termos lingüísticos utilizados (Muito Perto, Perto, Médio, Longe, Muito Longe).

Vamos supor que deseja-se efetuar a defuzzificação do valor lingüístico $d_{i,z} = \text{Médio}$, cujo valor da função de pertinência é $\mu_{\text{Médio}} = 0,8$. A partir desta informação devemos analisar qual o valor da função de pertinência do termo lingüístico anterior e posterior a *médio*. Desta forma, obteremos os seguintes valores: $\mu_{\text{Perto}} = 0,2$ e $\mu_{\text{Longe}} = 0$.

Portanto o valor de $d_{i,z}$, referente à defuzzificação do valor referenciado $d_{i,z}$ é:

$$(q_{\text{Perto}} * \mu_{\text{Perto}} + q_{\text{Médio}} * \mu_{\text{Médio}} + q_{\text{Longe}} * \mu_{\text{Longe}}) / (\mu_{\text{Perto}} + \mu_{\text{Médio}} + \mu_{\text{Longe}}).$$

$$(150 * 0.2 + 250 * 0.8 + 350 * 0) / (0.2 + 0.8 + 0) = 230$$

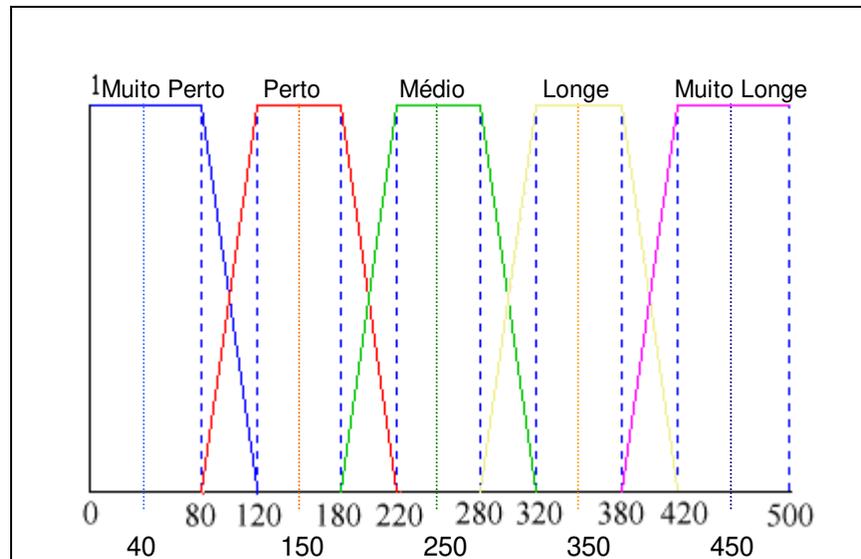


Figura 3.2 - Defuzzificação dos termos linguísticos de distância.

3.2 Linguagens Formais

A linguagem é um meio de comunicação, formada por um conjunto de palavras e de regras gramaticais que permitem combinar as palavras em sentenças sintaticamente corretas.

Uma linguagem é dita formal quando pode ser representada através de um sistema com sustentação matemática.

A Linguística Formal (Hopcroft & Ullman, 1979) compreende a representação da sintaxe (estrutura) e da semântica (significado) das sentenças de uma linguagem.

3.2.1 Alfabeto e Linguagens

Alfabeto, ou vocabulário, é um conjunto finito de símbolos, como por exemplo: dígitos, letras, letras gregas, etc (Hopcroft & Ullman, 1979).

Sentença ou palavra, definida sobre um alfabeto é qualquer seqüência finita de símbolos do alfabeto. Por exemplo:

alfabeto $V = \{0,1\}$
sentenças válidas = 001 e 010011

O tamanho, ou comprimento, de uma sentença é dado pelo número de símbolos que compõem a sentença. Por exemplo:

alfabeto $V = \{ a, b, c \}$

sentença $w = ab$

tamanho $|w| = 2$

Sentença vazia é a sentença que não contém símbolos, possuindo tamanho 0. É representada pelo símbolo ϵ (Hopcroft & Ullman, 1979).

Seja V um alfabeto, podemos representar por:

(XI) V^* = conjunto de todas as sentenças compostas de símbolos de V incluindo a sentença vazia,

e

(XII) $V^+ = V^* - \{\epsilon\}$

Exemplos:

$$V = \{ 0, 1 \}$$

$$V^* = \{ \epsilon, 0, 1, 00, 01, 10, 11, \dots \}$$

$$V^+ = \{ 0, 1, 00, 01, 10, 11, \dots \}$$

Linguagem é qualquer conjunto de sentenças sobre um alfabeto, ou seja, uma linguagem L sobre um alfabeto V é um subconjunto de V^* .

3.2.2 Gramáticas

Uma gramática serve para definir qual o subconjunto de sentenças que faz parte de uma determinada linguagem. É um dispositivo formal para especificar uma linguagem potencialmente infinita de uma forma finita.

Neste trabalho, representaremos linguagens a partir da notação formal de gramática.

3.2.2.1 Notação Formal de Gramática

Uma gramática G é definida por uma quádrupla (Hopcroft & Ullman, 1979):

(XIII) $G = (N, T, P, S)$ onde,

- a) **N**: é um conjunto finito de símbolos não-terminais;
- b) **T**: é um conjunto finito de símbolos terminais;
- c) **P**: é um conjunto finito de regras de produção, onde cada produção se encontra no formato $A \rightarrow \alpha$, onde **A** é um símbolo não-terminal e α é uma cadeia de símbolos em $(N \cup T)^*$;
- d) **S**: é um símbolo inicial da gramática.

Sendo que:

(XIV)
$$N \cap T = \emptyset$$
$$V = N \cup T$$
$$S \in N$$
$$P = \{ \alpha \rightarrow \beta \mid \alpha \in V^+ \text{ e } \beta \in V^* \}$$

Exemplo – Números Naturais:

$G = (N, T, P, I)$

$N = \{ I, D \}$

$T = \{ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \}$

$P = (I \rightarrow D, D \rightarrow 0, D \rightarrow 1, D \rightarrow 2, D \rightarrow 3, D \rightarrow 4, D \rightarrow 5, D \rightarrow 6, D \rightarrow 7, D \rightarrow 8, D \rightarrow 9, D \rightarrow DD)$

Segundo Rocha, (Rocha, 1997; Rocha et al., 1998) pode-se dizer que a aplicação de uma regra implica em três operações:

1. **Comparação**: o lado esquerdo da provável regra de reescrita é comparado com os símbolos da cadeia do lado direito a ser processada. Se esta comparação for bem sucedida, então:
2. **Reescrita**: a subcadeia selecionada do lado esquerdo é substituída pelo lado direito da regra de reescrita aceita. Finalmente:
3. **Avaliação**: valor booleano que informa se houve ou não a possibilidade de reescrita, ou seja, se o passo de reescrita foi executado.

3.2.2.2 Convenções

Para facilitar a compreensão das expressões, são adotadas sempre que possível as seguintes convenções (Hopcroft & Ullman, 1979):

- N** = { **A, B, C, D, E** } - Letras maiúsculas do início do alfabeto para símbolos não-terminais;
T = { **a, b, c, d, e** } - Letras minúsculas do início do alfabeto para símbolos terminais;
T* = { **u, v, ... z** } - Letras minúsculas do fim do alfabeto para seqüências de símbolos terminais;
(N ∪ T)* = { **α, β, γ, ...** } - Letras gregas para seqüências de símbolos não-terminais e terminais;
N ∪ T = { **X, Y, Z** } - Letras maiúsculas do fim do alfabeto para um símbolo terminal ou não-terminal.

3.2.2.3 Derivação

Derivação é uma operação de substituição efetuada de acordo com as regras de produção da gramática (Hopcroft & Ullman, 1979). Representa-se esta operação pelo símbolo \Rightarrow . Logo:

$$(XV) \quad \begin{array}{l} \gamma \alpha \delta \Rightarrow \gamma \beta \delta \\ \text{se } \alpha \rightarrow \beta \in P, \text{ e} \\ \gamma \text{ e } \delta \in (N \cup T)^* \end{array}$$

Uma seqüência de derivações, de zero ou mais passos, é representada por $*\Rightarrow$. Isto é:

$$(XVI) \quad \begin{array}{l} \alpha_1 * \Rightarrow \alpha_m \text{ se} \\ \alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m, \\ \alpha_1, \alpha_2, \dots, \alpha_m \in (N \cup T)^* \end{array}$$

3.2.2.4 Linguagem

A linguagem gerada por uma gramática **G** = (**N, T, P, S**) é o conjunto de todas as sentenças que podem ser geradas a partir do símbolo inicial desta gramática, através de derivações sucessivas (Hopcroft & Ullman, 1979).

$$(XVII) \quad L(G) = \{ w \mid w \in T^* \text{ e } S * \Rightarrow w \}$$

Exemplo:

$$\begin{array}{l} G = (N, T, P, S) \\ G = (\{S\}, \{0,1\}, P, S) \\ P = \{ S \rightarrow 0S1 \\ \quad S \rightarrow 01 \} \\ S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 03S13 \Rightarrow \dots 0_{n-1}S1_{n-1} \Rightarrow 0_n1_n \\ L(G) = \{ 0_n1_n \mid n \geq 1 \} \end{array}$$

3.2.3 Tipos de Gramáticas

Segundo a hierarquia de Chomsky, (Chomsky, 1965; Hopcroft & Ullman, 1979) as gramáticas podem ser de quatro tipos. Para $G=(N, T, P, S)$ e $V = N \cup T$, temos:

Tipo 0 - Gramáticas Irrestritas (GI)

A Gramática Irrestrita, é a maior família das gramáticas na hierarquia de Chomsky (Hopcroft & Ullman, 1979) e permite produções do tipo $\alpha \rightarrow \beta$, onde α e β são cadeias de caracteres arbitrárias e $\alpha \neq \epsilon$.

Ou seja, do lado esquerdo da produção pode haver uma seqüência de quaisquer símbolos, desde que, entre eles, haja um não-terminal. Do lado direito da produção pode haver qualquer seqüência de símbolos, inclusive a sentença vazia.

Tipo 1 - Gramáticas Sensíveis ao Contexto (GSC)

As Gramáticas Sensíveis ao Contexto requerem produções da forma $\alpha_1 A \alpha_2 \rightarrow \alpha_1 \beta \alpha_2$, onde α_1 e $\alpha_2 \in V^*$, $A \in N$ e $\beta \in V^+$ e o comprimento da sentença do lado esquerdo deve ser menor ou igual ao comprimento da sentença do lado direito da produção.

Ou seja, α_1 e α_2 são cadeias arbitrárias, incluindo a sentença vazia. A pertence ao conjunto dos símbolos não-terminais e β é uma cadeia não vazia, pertencente à união dos conjuntos dos símbolos terminais com os não-terminais. Informalmente, podemos dizer que: “ A pode ser substituído por β no contexto α_1, α_2 ”.

Tipo 2 - Gramáticas Livres de Contexto (GLC)

Em uma Gramática Livre de Contexto, cada produção é da forma $A \rightarrow \alpha$, onde $A \in N$ e $\alpha \in V^*$. No lado esquerdo da produção, deve existir apenas um símbolo, o qual deve ser um não-terminal.

Tipo 3 - Gramáticas Regulares (GR)

Se todas as produções de uma **GLC** são da forma $A \rightarrow wB$ ou $A \rightarrow w$, onde **A** e **B** são símbolos não-terminais e **w** é uma cadeia de caracteres formada por símbolos terminais podendo ser uma sentença vazia, então esta gramática é chamada de linear-direita. Se todas as produções são da forma $A \rightarrow Bw$ ou $A \rightarrow w$, esta gramática é chamada de linear-esquerda. Uma gramática linear-direita ou esquerda é chamada gramática regular.

A Figura 3.3 mostra a relação entre os tipos de gramáticas.

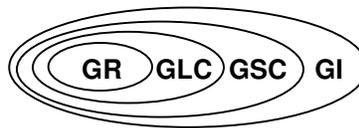


Figura 3.3 - Relação entre os tipos de gramáticas.

3.2.3.1 Linguagens geradas por gramáticas

Conforme o tipo da gramática que dá origem a uma linguagem, esta se classifica em:

- **LSC** - Linguagem Sensível ao Contexto;
- **LLC** - Linguagem Livre de Contexto;
- **LR** - Linguagem Regular.

3.2.4 Gramáticas Livres de Contexto (GLC)

As Gramáticas Livres de Contexto têm grande importância dentro do estudo das Linguagens Formais, pois através delas pode ser descrita a maior parte das construções sintáticas das linguagens de programação. A seguir, um exemplo de Gramática Livre de Contexto:

$G = (N, T, P, S)$

$N = \{S, A, B\}$

$T = \{a, b\}$

$P = \{ S \rightarrow aB \quad A \rightarrow bAA$
 $S \rightarrow bA \quad B \rightarrow b$
 $A \rightarrow a \quad B \rightarrow bS$
 $A \rightarrow aS \quad B \rightarrow aBB \}$

3.3 Linguagem Formal Fuzzy

A definição de uma Linguagem Formal e uma gramática **G**, descrita no item 3.2 - Linguagens Formais (Chomsky, 1965, Hopcroft & Ullman, 1979, Simpser, 1997), foi ampliada para que pudesse representar também o conceito de conjuntos nebulosos.

A teoria dos conjuntos nebulosos evoluiu o conceito de incerteza, que foi incorporado às Linguagens Formais, gerando desta forma, as Linguagens Formais Fuzzy ou Nebulosas.

Uma das principais mudanças foi a inclusão do conceito de incerteza na reescrita (Mizumoto et al., 1973, Negoita & Ralescu, 1975), comparação e avaliação (Rocha, 1992, Rocha, 1997).

Dessa forma, pode-se dizer que uma gramática **G** é uma estrutura definida como (Mizumoto et al., 1973; Rocha, 1992; Rocha, 1997):

$$(XVIII) \quad \mathbf{G} = (\mathbf{N}, \mathbf{T}, \mathbf{P}, \mathbf{S}, \mu)$$

onde:

- a) **N**: é um conjunto finito de símbolos não-terminais;
- b) **T**: é um conjunto finito de símbolos terminais;
- c) **P**: é um conjunto finito de regras de produção, onde cada produção se encontra na seguinte forma $r: \alpha \rightarrow \beta(\mu_r)$, onde α é um símbolo não-terminal e β é uma cadeia de símbolos em $(\mathbf{N} \cup \mathbf{T})^*$.
O grau associado a produção r é denotado por μ_r .
- d) **S**: é um símbolo inicial da gramática.
- e) μ : é uma função de pertinência de $\mu \rightarrow [0,1]$ que representa o grau de aplicação de uma produção.

Também podemos utilizar as definições feitas em (XI) e (XII), para representar um alfabeto $\mathbf{V} = \mathbf{N} \cup \mathbf{T}$:

V* = conjunto de todas as sentenças compostas de símbolos de **V** incluindo a sentença vazia.

$$\mathbf{V}^+ = \mathbf{V}^* - \{\epsilon\}$$

A seqüência de derivação $\alpha_1 \Rightarrow \alpha_m$ da cadeia $\alpha_m \in \mathbf{V}^*$ de **G** é o conjunto ordenado de reescritas requeridas para transformar o símbolo inicial $\alpha_1 \in \mathbf{S}$ em α_m . Em outras palavras:

$$\alpha_1 \stackrel{*}{\Rightarrow} \alpha_m \text{ se}$$

$$(XIX) \quad \alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m, \\ \alpha_1, \alpha_2, \dots, \alpha_m \in (\mathbf{N} \cup \mathbf{T})^*$$

Uma linguagem formal \mathbf{L} é definida como um subconjunto de cadeias aceitas por uma gramática \mathbf{G} . As cadeias α_m geradas por \mathbf{G} e aceitas como pertencentes a \mathbf{L} são chamadas **fórmulas bem formadas (fbf)** de \mathbf{L} , de acordo com \mathbf{G} . Uma cadeia \mathbf{s}_j produzida por \mathbf{G} é uma **fbf** de \mathbf{L} , se ela pertencer a \mathbf{T} . Em outras palavras, a cadeia α_m aceita pela linguagem $\mathbf{L}(\mathbf{G})$, suportada por \mathbf{G} , são aquelas **fbf**(α_1, α_m) obtidas em:

$$(XX) \quad \mathbf{fbf}(\alpha_1, \alpha_m) = \alpha_1 \stackrel{*}{\Rightarrow} \alpha_m = \alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m, \alpha_m \in \mathbf{T}, \alpha_1 \in \mathbf{S}$$

O processamento de qualquer sentença da seqüência de derivação $\mathbf{s}_i \stackrel{*}{\Rightarrow} \mathbf{s}_m$ é um conjunto **seqüencialmente ordenado** de operações de reescrita, cada uma envolvendo os seguintes passos:

1) **Comparação**: o lado esquerdo \mathbf{s}_i da provável regra de reescrita é comparado com os símbolos da cadeia \mathbf{s}_j a ser processada. Se esta comparação for bem sucedida, ou seja $\mu(\mathbf{s}_i, \mathbf{s}_j) > \theta$, então:

2) **Reescrita**: a subcadeia selecionada em \mathbf{s}_i é substituída pelo lado direito \mathbf{s}_j da regra de reescrita aceita. Finalmente:

3) **Avaliação**: a pertinência da cadeia derivada \mathbf{s}_j a \mathbf{T} é calculada como uma medida definida no intervalo fechado $[0,1]$.

Agora, considerando as seguintes definições (Rocha, 1997, Rocha et al., 1998):

a) O **grau de similaridade** (comparação) $\mu(\mathbf{s}_i, \mathbf{s}_j)$ de duas cadeias $\mathbf{s}_i, \mathbf{s}_j$ é uma medida do espaço cartesiano $(\mathbf{V}^* \times \mathbf{V}^*)$ para o intervalo fechado $[0,1]$. Em outras palavras:

$$(XXI) \quad \mu : \mathbf{V}^* \times \mathbf{V}^* \rightarrow [0, 1] \\ \mu(\mathbf{s}_i, \mathbf{s}_j) \rightarrow 1 \text{ se } \mathbf{s}_i \text{ tende a ser igual a } \mathbf{s}_j; \quad \mu(\mathbf{s}_i, \mathbf{s}_j) \rightarrow 0 \text{ caso contrário.}$$

b) O **grau de avaliação** de \mathbf{s}_j como uma **fbf** de \mathbf{L} é calculado através do grau de similaridade $\mu(\mathbf{s}_j, \mathbf{s}_i)$ de \mathbf{s}_j para com a cadeia $\mathbf{s}_i \in \mathbf{T}^*$. Sendo assim:

$$(XXII) \quad \mu(\mathbf{s}_j, \mathbf{s}_i) \rightarrow [0, 1] \\ \mu(\mathbf{s}_j, \mathbf{s}_i) \rightarrow 1 \text{ se } \mathbf{s}_j \text{ tende a ser uma } \mathbf{fbf}, \text{ já que tende a ser igual a } \mathbf{s}_i.$$

As restrições em (XVIII), (XXI) e (XXII) definem **L** como uma linguagem nebulosa (Mizumoto et al., 1973; Negoita & Ralescu, 1975; Rocha, 1997; Rocha et al., 1980).

Agora, se:

a) $\mu(s_i, s_j)$ em (XXI) é tomada como uma função binária de valor 1 ou 0, isto é,

$$(XXIII) \quad \mu: V^* \times V^* \rightarrow [0,1] \\ \mu(s_i, s_j) = 1 \text{ se } s_i \text{ igual a } s_j ; \quad \mu(s_i, s_j) = 0 \text{ caso contrário.}$$

b) $\mu(s_j, T^*)$ em (XXII) é tomada como sendo similar a uma função identidade, isto é:

$$(XXIV) \quad \mu(s_j, T^*) = 1 \text{ se } s_j \in T^* , \text{ caso contrário } \mu(s_j, T^*) = 0$$

então, **L** torna-se uma linguagem clássica do tipo discutido por Chomsky e outros (Chomsky, 1965; Searls, 1993).

3.4 Sistema Visual Humano

Diversos trabalhos mostram a teoria de um sistema visual artificial, a partir do conhecimento sobre a visão biológica (Rocha, 1997; Serapião et al., 1996; Alegre et al. 1993a; Alegre et al. 1993b).

Rocha (Rocha, 1999) afirma que o cérebro utiliza distintos receptores para medir diferentes tipos de variações energéticas, coletando, deste modo, dados sobre a dinâmica do ambiente em que vive. Os dados sensoriais são coletados pelos receptores sensoriais e processados por neurônios que se distribuem por várias partes do cérebro. O conjunto de neurônios que se especializam para o processamento dos dados sensoriais formam um sistema sensorial.

As ondas luminosas emitidas pelos objetos que vemos no mundo que nos cerca criam uma imagem deste objeto sobre a retina e estimulam dois tipos diferentes de receptores: os bastonetes e os cones. Os cones são sensíveis a três cores: verde, vermelho e azul, gerando assim três imagens, enquanto que os bastonetes são ativados por ondas luminosas de qualquer frequência, medindo assim a luminância da cena observada e gerando uma imagem em tons de cinza (Rocha, 1999).

A retina realiza um processamento complexo da imagem visual e gera tipos distintos de informação para as outras áreas cerebrais (Rocha, 1999):

1. três descrições nas cores básicas verde, vermelho e azul;
2. uma descrição da luminância da cena;
3. uma descrição dos contrastes de luminosidade;
4. uma descrição dos movimentos dos elementos da cena.



Figura 3.4 - Etapas da obtenção de uma imagem pelo sistema visual humano.

O papel dos neurônios localizados em diversas áreas cerebrais é analisar características particulares em cada uma dessas diferentes imagens geradas pela retina e recombina-las para reconstruir a cena original, da maneira como iremos finalmente enxergá-la. Temos, então, um processo que inicialmente decompõe a cena e analisa características especiais da imagem, e depois remonta a realidade a partir desta análise primária.

3.4.1 Reconhecimento de propriedades básicas das imagens

Inicialmente, as informações provenientes da retina ativam as áreas visuais primárias, localizadas no lobo occipital, cuja função é extrair características básicas de imagem, tais como direções predominantes de linhas; identificação de ângulos (Figura 3.5); composição de cores (Figura 3.6); análise de direções e velocidades de movimentação (Figura 3.7), etc. O papel destas áreas no processamento visual é, portanto, decompor a imagem em um conjunto de características básicas, que serão utilizadas no reconhecimento dos elementos que compõem a cena a ser analisada (Rocha, 1999).

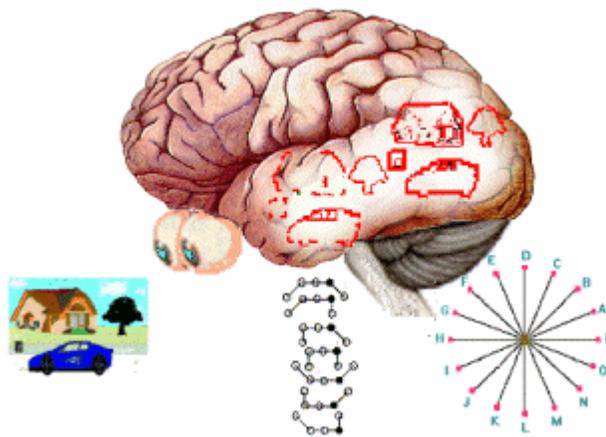


Figura 3.5 - A análise cortical de linhas e ângulo.

Na seqüência, o resultado do processamento nas áreas visuais secundárias é transferido para áreas de associação, distribuídas nos lobos temporal e parietal. O papel destas áreas de associação é reconhecer os principais elementos que compõem a cena, a partir de suas características básicas. Um quadrado é reconhecido por ser composto por linhas de tamanho semelhante e que formam ângulos de aproximadamente 90°. Uma casa será identificada por ser composta por linhas de direções específicas, que formam ângulos característicos, etc. Cada elemento é, portanto, reconhecido pelas relações específicas partilhadas por um conjunto de características definidas. Quando se aprende a reconhecer um objeto, se aprende a identificar este conjunto de características específicas e as relações estabelecidas entre elas.

Enquanto os objetos principais de uma cena são reconhecidos pelas suas formas, outras áreas visuais fazem uma análise da distribuição das cores nos campo visuais (Figura 3.4) ou dos movimentos aí registrados (Figura 3.7). Estas análises ocorrem em paralelo e dissociadas das análises de formas. A unificação de todas estas informações para recompor a cena observada é tarefa de áreas de associação secundárias, localizadas, principalmente, no córtex parietal.

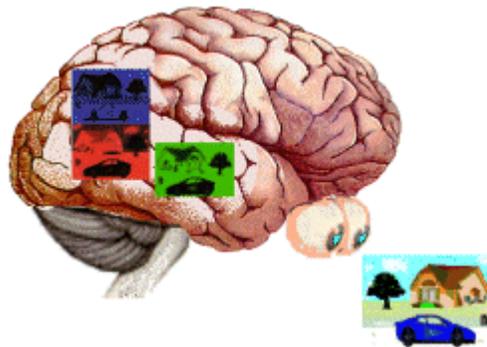


Figura 3.6 - A análise cortical de cores.

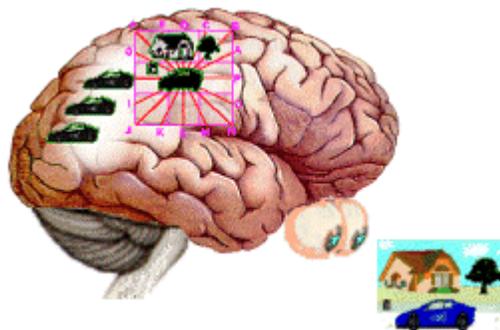


Figura 3.7 - A análise cortical do movimento e localização espacial.

3.4.2 Recomposição de uma cena

A segunda etapa do processamento sensorial envolve a recomposição da cena final a partir dos elementos individualmente reconhecidos, e das relações partilhadas por eles na imagem observada e identificadas pelo cérebro.

É postulado que a identificação dos objetos é um trabalho preferencial do hemisfério esquerdo, enquanto que a análise das relações espaciais fica mais a cargo do hemisfério cerebral direito (Rocha, 1999), como é ilustrado na Figura 3.8. Esta análise espacial envolve também a composição e análise da coerência dos movimentos associados a cada um dos objetos que compõem a cena visualizada.

Supõe-se, assim, que o hemisfério direito tem um papel muito importante no processamento visual, uma vez que a recomposição espacial das relações partilhadas pelos objetos da cena é um processamento muito importante para a criação da imagem visualizada da cena original.

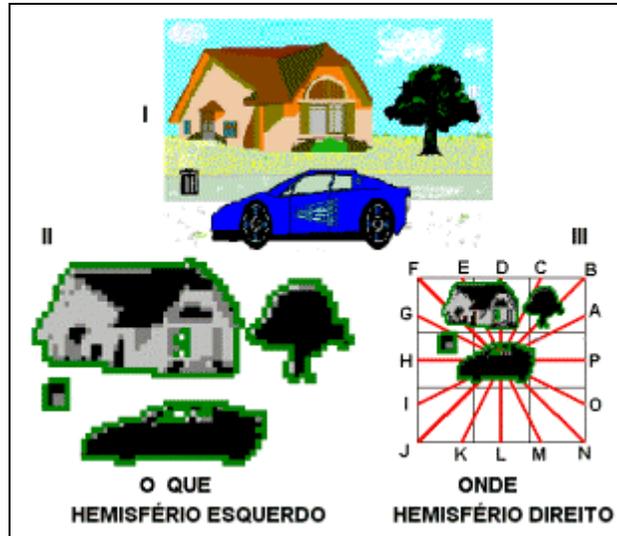


Figura 3.8 - A especialização hemisférica.

3.4.3 Um Sistema Visual Formal

Conforme proposto por Rocha (Rocha, 1997) e Serapião (Serapião et al., 1996), a teoria de linguagens formais nebulosas pode ser utilizada para simular o funcionamento do sistema visual humano. O processamento no nível da retina pode ser simulado por vários tipos de filtros e algoritmos para extração de contorno, enquanto que filtros de segmentação podem simular circuitos pré-corticais (Figura 3.9).

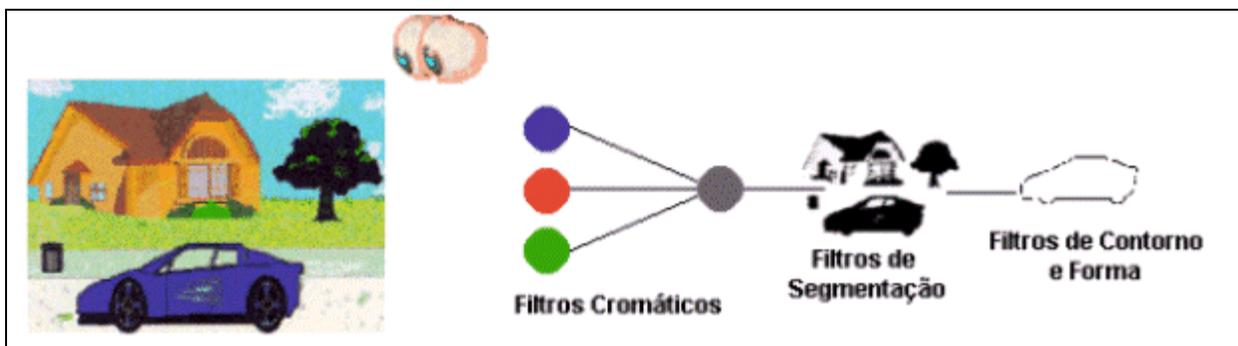


Figura 3.9 - Organização computacional dos filtros.

Podemos dizer então que uma imagem refere-se a uma função de intensidade luminosa bidimensional, denotada por $f(x,y)$, em que o valor ou amplitude de f nas coordenadas espaciais (x,y) dá a intensidade (brilho) da imagem naquele ponto. Como a luz é uma forma de energia, $f(x,y)$ deve ser positiva e finita, isto é:

(XXV)
$$0 < f(x, y) < \infty$$

As imagens que as pessoas percebem em atividades visuais corriqueiras consistem de luz refletida dos objetos. A natureza básica de $f(x,y)$ pode ser caracterizada por dois componentes:

- (1) a quantidade de luz incidida na cena sendo observada;
- (2) a quantidade de luz refletida pelo objetos na cena.

Apropriadamente, esses componentes são chamados de iluminação e reflectância, respectivamente, e são representados por $i(x,y)$ e $r(x,y)$. O produto das funções $i(x,y)$ e $r(x,y)$ resulta $f(x,y)$.

$$f(x, y) = i(x,y) r(x,y)$$

(XXVI)
$$\begin{array}{ll} \text{Onde,} & 0 < i(x,y) < \infty \\ \text{e,} & 0 < r(x,y) < 1 \end{array}$$

A equação acima indica que a reflectância (r) é limitada entre 0 (absorção total) e 1 (reflectância total). A natureza de $i(x,y)$ é determinada pela fonte de luz, e $r(x,y)$ é determinada pelas características dos objetos na cena.

Denominamos a intensidade de uma imagem monocromática f nas coordenadas (x,y) de nível de cinza (lu) da imagem naquele ponto. Como mostra as equações acima, fica evidente que $lu(x,y)$ fica restrito ao intervalo.

(XXVII)
$$LU_{\min} \leq lu(x,y) \leq LU_{\max}$$

Em teoria, a única restrição sobre LU_{\min} é que seja um valor positivo e sobre LU_{\max} é que seja maior que LU_{\min} .

O intervalo $[LU_{\min} , LU_{\max}]$ é denominado escala de cinza. A prática comum é deslocar este intervalo para $[0 , LU]$, onde $lu(x,y) = 0$ é considerado preto e $lu(x,y) = LU$ é considerado branco. Todos os valores intermediários são tons de cinza variando continuamente entre o branco e o preto.

Supondo que seja dado um dispositivo de sensoriamento S para observar um conjunto de objetos O_i , pode-se definir a imagem I_i gerada por S para a observação do objeto O_i como:

(XXVIII)
$$I_i = \{ x_j , y_j , m_j \}_{j=1 \text{ até } k}$$

onde m_j é o valor da medida numérica realizada por S na posição (x_j, y_j) sobre O_i , e k é o número de imagens I . O conjunto de posições que podem ser sensoriadas por S definem um reticulado L . O sistema visual mostrado na Figura 3.10 é composto de um dispositivo **Lum** para medida da luminância sobre L_1 . Esse sistema pode ser utilizado para gerar imagens visuais I_i de objetos O_i focados sobre L_1 , isto é:

$$I_i = \{ x_j, y_j, lm_j \}_{j=1 \text{ até } k}$$

onde lm_j é a medida de luminância realizada na posição (x_j, y_j) .

Suponha um esquema como o **Sensor** (Serapião et al., 1996) que é um sistema hierárquico de sensoriamento inteligente que utiliza agentes primários para o reconhecimento de características elementares de objetos sensoriados por um dispositivo visual V , e outros agentes de maior grau hierárquico para o reconhecimento progressivo de características mais complexas dos objetos O_i observados por V . Uma das características dos agentes do sistema **Sensor** é sua capacidade de utilizar linguagens formais definidas por gramáticas G_j para recodificar as informações numéricas contidas em I_i .

No exemplo da Figura 3.10, descrito em Rocha, 1997, o agente primário **Cont** do **Sensor** recodifica a imagem I_i gerada por V sobre um objeto O_i em uma sentença s_1 de caracteres $d^{(j)}$ pertencentes a um dicionário D_1 { $p = \text{preto}$, $b = \text{branco}$ } e suportada por uma gramática regular G_1 , isto é:

$$(XXIX) \quad s_1 = d^{(k)}$$

tal que:

$$(XXX) \quad d^{(j)} = p \text{ se } lm_j < \alpha, \text{ caso contrário } d^{(j)} = b, \text{ na posição } x_j, y_j$$

onde, $d^{(j)}$ é o j -ésimo caractere de s_1 e k é igual ao número de posições de sensoriamento em L_1 .

O agente **Bor** na Figura 3.10, reescreve a sentença s_1 em outra sentença s_2 suportada por uma gramática regular G_2 definida sobre o seguinte dicionário $D_2 = \{n, f\}$ de acordo com as regras:

$$(XXXI) \quad pb \rightarrow n, bp \rightarrow n, bb \rightarrow f, pp \rightarrow f$$

Em outras palavras, o agente **Bor** gera uma sentença s_2 que descreve a borda do objeto O_i sensoriado por V , sempre que s_2 contiver uma subcadeia:

$$(XXXII) \quad B_2 = d_2^k \dots d_2^m$$

tal que, qualquer $d_2^k = n$ para qualquer $d_2^k \in B_2$ e k e m estão associados a pontos vizinhos em L_1 .

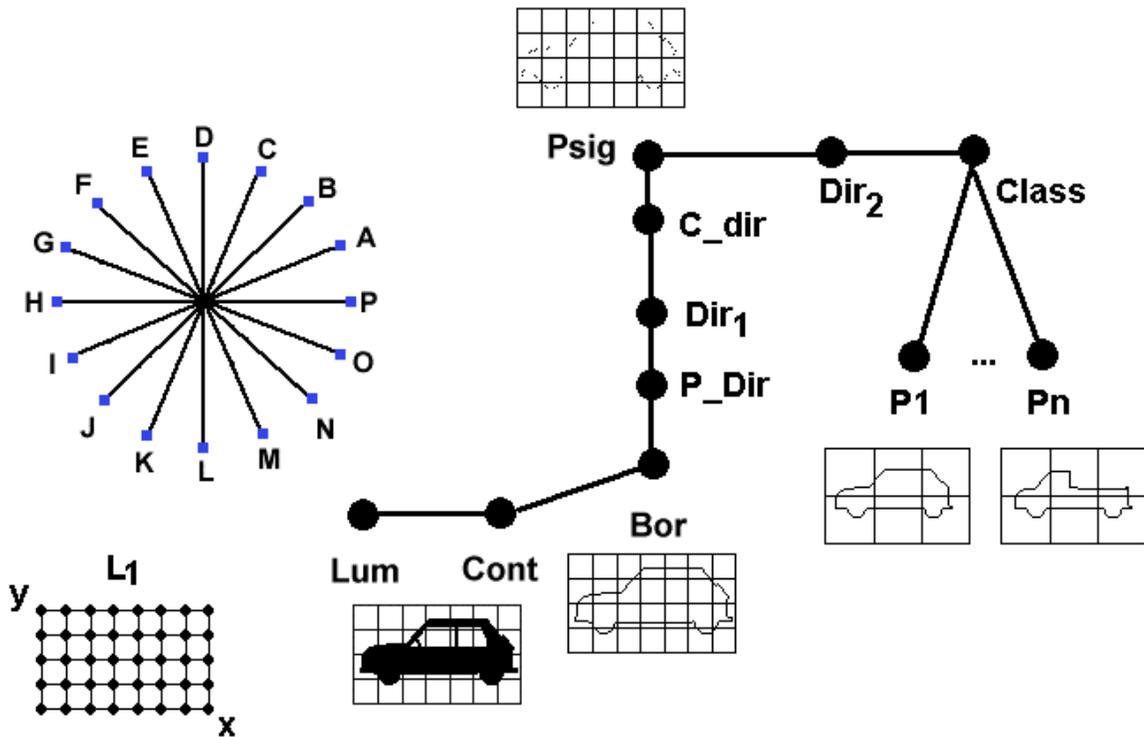


Figura 3.10 - Exemplo de um sistema de inteligência artificial, composto por objetos de reconhecimento de elementos, analisando suas características e complexidades. Estes agentes são desenvolvidos para descrever o comportamento dos neurônios do sistema visual natural e segue a mesma hierarquia organizacional das áreas do sistema visual natural.

A localização l_{ij} entre duas posições l_i, l_j em L_1 pode ser definida como:

(XXXIII)
$$l_{ij} = \{ l_x, l_y \}$$

$$l_x = x_i - x_j, l_y = y_i - y_j$$

O agente P_Dir na Figura 3.10, gera uma sentença s_3 que descreve as direções entre dois caracteres consecutivos de uma borda B_2 , utilizando o dicionário D_3 , onde:

(XXXIV) $D_3 = \{ p = \text{positivo}, n = \text{negativo}, z = \text{zero}, a = \text{vazio}, b = \text{pequeno}, c = \text{médio}, d = \text{grande}, e = \text{enorme} \}$

tal que,

(XXXV)

Negativo	l_x ou $l_y < 0$
Positivo	l_x ou $l_y > 0$
Zero	l_x ou $l_y = 0$
Vazio	$l_x / l_y < \alpha_1$
Pequeno	$\alpha_1 \leq l_x / l_y < \alpha_2$
Médio	$\alpha_2 \leq l_x / l_y < \alpha_3$
Grande	$\alpha_3 \leq l_x / l_y < \alpha_4$
Enorme	$l_x / l_y \geq \alpha_4$

onde s_3 é uma sentença suportada por uma gramática G_3 .

$$(XXXVI) \quad s_3 = (c_{x1} c_{y1} c_2)^*$$

tal que:

$$(XXXVII) \quad c_{x1}, c_{y1} \in \{p, n, z\}, c_2 \in \{b, c, d, e\}$$

Em outras palavras, s_3 descreve as direções entre elementos consecutivos da borda B_2 através de uma gramática, onde c_2 classifica um ângulo e c_{x1} e c_{y1} identificam o seu quadrante.

O agente Dir_1 reescreve s_3 em uma outra cadeia s_4 suportada pela gramática G_4 , utilizando o seguinte dicionário:

$$(XXXVIII) \quad D_4 = \{A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P\},$$

de acordo com as seguintes regras:

$$(XXXIX) \quad \left\{ \begin{array}{llll} p^*a \rightarrow P & ppb \rightarrow A & ppc \rightarrow B & ppd \rightarrow C \\ *pe \rightarrow D & npd \rightarrow E & npc \rightarrow F & npb \rightarrow G \\ n^*a \rightarrow H & nnb \rightarrow I & nnc \rightarrow J & nnd \rightarrow K \\ *ne \rightarrow L & pnd \rightarrow M & pnc \rightarrow N & pnb \rightarrow O \end{array} \right\}$$

onde “*” significa “qualquer direção”.

Por exemplo, a cadeia gerada pelo agente Dir_1 , que descreve o contorno do carro na Figura 3.10:

HDDCCPCPPPPPCCCCPPPPPPPPPMPMCMLLKKHEEEHHHHHHHHHHHHHHKKKEHEHHHD

Na seqüência, o agente C_Dir reescreve a cadeia s_4 em uma cadeia s_5 composta por símbolos deste dicionário:

$$(XL) D_5 = \{z = \text{Sem mudança}, c = \text{Mudança no sentido anti-horário}, d = \text{Mudança no sentido horário}\},$$

descrevendo as mudanças de direções entre os pontos da figura, associadas com dois símbolos consecutivos de s_4 . Assim, s_5 descreve as mudanças de direções da borda da figura do carro.

O agente $Psig$ pesquisa a existência das subcadeias: **zcc, dcc, zdd, cdd, czc, dzd, zzc, zzd, zcz, e zdz**, em s_5 para gerar um vetor v_6 contendo coordenadas **xy** da localização das subcadeias. Este vetor descreve os pontos da borda **Psig** (pontos significativos) onde as mudanças de direções ocorreram.

O agente **Psig** usa a seguinte gramática G_5 sobre o dicionário D_5 , para gerar uma cadeia s_6 de acordo com as seguintes regras:

$$(XLI) \quad \left[\begin{array}{cccc} zcc \rightarrow S & dcc \rightarrow S & zdd \rightarrow S & cdd \rightarrow S \\ czc \rightarrow S & dzd \rightarrow S & zzc \rightarrow S & zzd \rightarrow S \\ zcz \rightarrow S & zdz \rightarrow S & & \end{array} \right]$$

$$\text{Senão} \rightarrow N,$$

produzindo uma cadeia do tipo:

$$(XLII) \quad s_6 = \{ \text{SNN.....} \}$$

Agora, o agente **Dir₂** (Figura 3.10) recodifica s_6 em uma cadeia s_7 descrevendo as principais direções dentre os pontos **Psig** da borda B_2 . No exemplo abaixo, somente os pontos significativos do objeto carro da Figura 3.10.

$$(XLIII) \quad s_7 = \{ \text{HDDBBPCCPPMMCMLKKFEHHJKFFHD} \}$$

A partir da informação gerada por s_7 , podemos agora armazenar esta informação como sendo um modelo de características de um objeto carro que é inserido em uma base de dados de conhecimento.

Com esta definição anteriormente descrita, novas imagens podem ser classificadas e seus objetos podem ser confrontados com a base de dados de conhecimento, a fim de retirar informações necessárias para a identificação de objetos.

3.5 Resumo do Capítulo

Procurou-se, neste capítulo, apresentar de forma sintética algumas noções teóricas necessárias para o entendimento deste trabalho.

O próximo capítulo, tratará dos procedimentos utilizados na pesquisa, tais como o grupo de voluntários selecionados, as atividades que estes voluntários desenvolveram e os softwares criados para que pudesse ser possível a realização deste trabalho.

4 Estudo Experimental

Um grupo de voluntários criou dois tipos de cenas: cena livre C_L e cena orientada C_O , utilizando um software desenvolvido especificamente para esse fim, e posteriormente fizeram uma descrição escrita I_k das cenas geradas.

Este software possui uma figura de fundo e uma série de objetos (elementos pictóricos) O_i que podem ser inseridos neste fundo, permitindo assim a construção das cenas c_k . Após a criação das cenas, o software deve requisitar uma descrição escrita I_k sobre o cenário gerado. O software deve possuir também a opção de tocar um som no início da aplicação. Este som pode conter tanto uma instrução, como a narrativa de uma história.

Dados os conjuntos C_L , C_O formados pelas cenas livres ou orientadas c_k , geradas por voluntários k , calculou-se, respectivamente, as cenas prototípicas médias \bar{c}_L e \bar{c}_O , visando com isso, representar o consenso da idéia dos cenários montados pelos voluntários. Para isto, será aplicada a técnica da TCP, proposta por Zadeh (Zadeh, 2001).

Para avaliar as cenas \bar{c}_L e \bar{c}_O geradas, serão criadas também cenas médias aplicando um cálculo de promediação convencional, que gerará cenas médias livres \bar{c}_{mL} e orientadas \bar{c}_{mO} . Os conceitos de cenas médias e cenas prototípicas médias, serão detalhadas no próximo capítulo.

Sendo L , o conjunto formado pelas descrições I_k das cenas c_k escritas por voluntários k , foi utilizado o sistema Jargão (Rocha, 1992; Rocha et al., 1992; Guilherme, 1996) para analisar este conjunto L de descrições com o intuito de obter descrições de consenso \bar{I} . Estudou-se também a relação destas descrições \bar{I} com os elementos pictóricos O_i utilizados na composição das cenas.

4.1 Grupo Experimental

Selecionou-se um grupo de 23 adultos, com nível de escolaridade médio ou superior que desenvolveu as seguintes tarefas:

- Ouviram um texto gravado de uma história de humor escrita por Luis Fernando Veríssimo, apresentado na Tabela 4.1;
- Ouviram novamente o trecho inicial da história, marcado em azul na Tabela 4.1 e compuseram uma cena sobre o trecho ouvido. Essa atividade recebeu o nome de “Abertura”;

- Compuseram uma cena sobre o episódio da história considerado mais engraçado;
- Desenvolveram uma cena sobre o tema natalino Fábrica de Brinquedos.
- Ao término da composição das cenas do episódio mais engraçado e da fábrica de brinquedos, os voluntários fizeram uma descrição escrita das cenas montadas.

O RECITAL.

Uma boa maneira de começar um conto é imaginar uma situação rigidamente formal – digamos, um recital de quarteto de cordas, todos de fraque impecavelmente engomados. Estão todos prontos, vai começar o recital. Nisso...

Nisso, o quê? Qual é a coisa mais insólita que pode ocorrer num recital de um quarteto de cordas? Passar uma manada de zebus pelo palco, por detrás deles? Não. É preciso instalar-se no acontecimento, como a semente da confusão, uma pequena incongruência. Algo que crie apenas um mal estar, de início, e chegue lentamente ao caos. Entra no palco um homem carregando uma tuba.

Há um murmúrio na platéia. O quê é aquilo? O homem entra com sua tuba dos bastidores. Posta-se ao lado do violoncelista, o qual tinha um grande bigode ruivo, isso pode ser importante. Depois de alguns instantes em que a tensão no ar é como uma corda de violino esticada ao máximo, o primeiro violinista fala:

- Por favor,...
- O quê? Vai dizer que eu não posso ficar aqui?
- O quê o senhor quer?
- Quero tocar, ora. Podem começar que eu acompanho.
- Por favor, isso é um recital de cordas de Mozart, não tem nenhuma parte para tuba.
- Eu improviso. Vocês começam e eu faço um-pá-pá.

Risos na platéia. De onde terá surgido aquele homem. Nem de fraque ele está. Segundo algumas versões ele vestia uma camisa do Vasco, e calçava um chinelo de dedos. O violinista ameaça chamar a polícia. Então o tubista fala:

- Se alguém se aproximar de mim eu toco pof.

A perspectiva de se ouvir um pof naquele recinto paralisa a todos. O tubista torna-se ameaçador:

- Vocês pensam que estão no século dezoito. Já houve 17 revoluções populares depois de Mozart. Eu vou confiscar suas partituras em nome do povo.

Depois num tom sedutor para a violoncelista:

- Eu represento seus sonhos secretos. Durante o Mozart é em mim que você pensa, na minha barriga e na minha tuba fálica, você quer ser violada por mim durante um alegre assai, confessa...

Finalmente, desafiador fala ao violoncelista:

- Esse bigode ruivo é meu. É o mesmo que eu usava em 68, devolve.

Os dois se atacam, o resto do quarteto entra na briga. A platéia grita e pula, é o caos. Um dos resultados da briga é que agora o bigode do violoncelista está na violoncelista. Vendo-a assim, o tubista para de morder a perna do violinista e grita:

- Mamãe!

Nisso entra no palco uma manada de zebus.

Tabela 4.1 - Texto humorístico de Luis Fernando Veríssimo – O Recital.

A quantidade de 23 voluntários não foi definida premeditadamente. Apenas foi a quantidade máxima de voluntários obtidos que se enquadraram nos requisitos definidos (nível de escolaridade médio ou superior).

A escolha dos temas das atividades (tema natalino e o texto humorístico "O Recital", de Luis Fernando Veríssimo) deve-se ao fato dos mesmos já terem sido empregados pelo grupo de pesquisa ao qual faço parte, para estudar a influência da ilustração dos episódios sobre a decodificação verbal e a atividade cerebral (Rocha et al., 2002), já que foi possível reutilizar o som, já gravado, da narração da história e as ilustrações feitas por um desenhista profissional.

Não foi estipulado um limite de tempo para que os voluntários realizassem a montagem das cenas e as descrições escritas das cenas montadas. Alguns voluntários se recusaram a fazer a descrição escrita da cena montada.

A figura de fundo e os elementos pictóricos utilizados na montagem das cenas c_k para as atividades "Abertura" e "Cena Mais Engraçada" são mostrados respectivamente na Figura 4.1 e Figura 4.2.

Na atividade "Abertura", pediu-se aos voluntários que utilizassem os elementos pictóricos O_i disponibilizados e criassem uma cena c_k que dissesse respeito ao trecho da história ouvido no início da tarefa (trecho em azul na Tabela 4.1).

Já na atividade "Cena Mais Engraçada", pediu-se aos voluntários que utilizassem os O_i disponibilizados e criassem uma cena c_k que demonstrasse o que eles haviam achado mais engraçado no conto "O Recital" (Tabela 4.1).



Figura 4.1 - Elementos Pictóricos e Figura de Fundo da atividade “Abertura”.

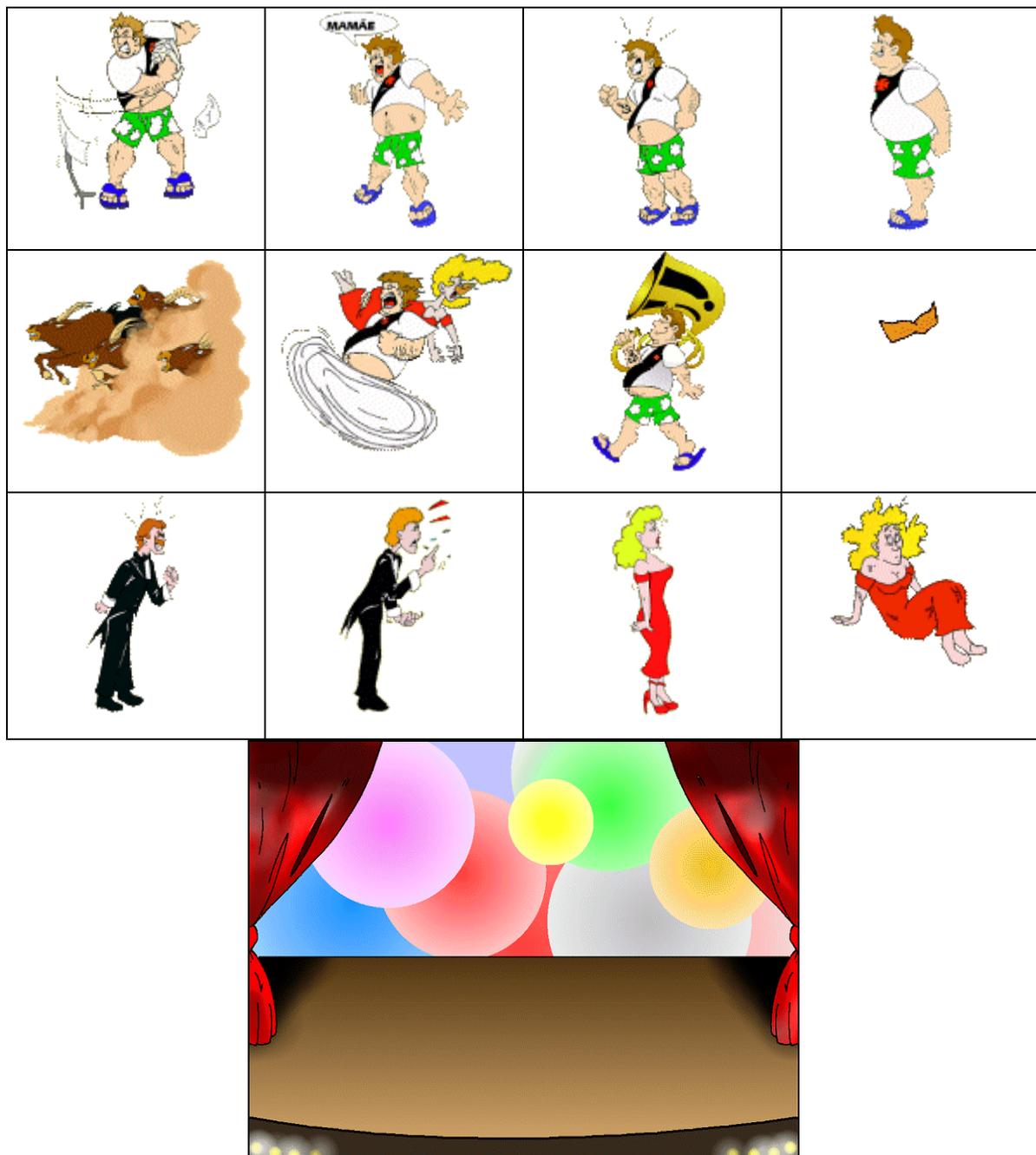


Figura 4.2 - Elementos Pictóricos e Figura de Fundo da atividade “Cena Mais Engraçada”.

A Figura 4.3 mostra a figura de fundo e os elementos pictóricos O_i disponibilizados aos voluntários para a composição de uma cena livre sobre o tema “Fábrica de Brinquedos do Papai Noel no Pólo Norte”.

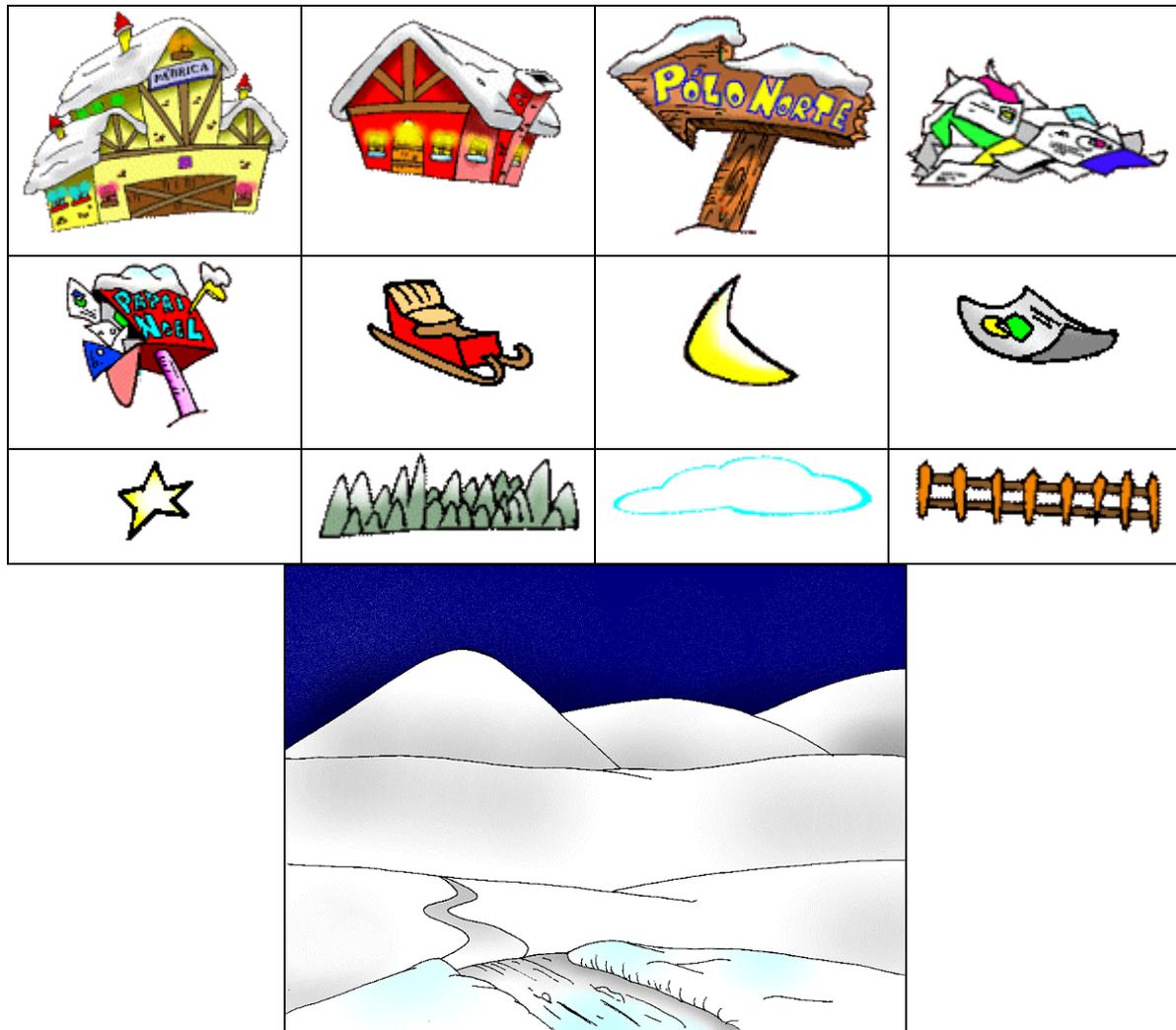


Figura 4.3 - Elementos Pictóricos e figura de fundo da atividade Fábrica de Brinquedos.

4.2 Ferramentas computacionais

Para a realização deste trabalho, foi necessário, portanto, o desenvolvimento de dois softwares:

- Um editor de cenários que possua ferramentas necessárias para a realização do estudo proposto, sendo que este deve ser orientado a objetos, e permitir tanto a montagem de cenas livres C_L , como cenas orientadas C_O a partir de uma narração tocada antes da montagem da

cena. É necessário também um módulo que guarde a descrição escrita da cena montada pelos voluntários.

- Um visualizador e manipulador das cenas c_k obtidas pelo grupo de voluntários, que além de permitir visualizar e navegar pelas cenas, aplique também determinados algoritmos para tentar gerar cenas prototípicas médias, tanto para cenas livres \bar{c}_L , quanto para cenas orientadas \bar{c}_O . Este software deve ser integrado ao sistema Jargão (Rocha, 1992; Rocha et al., 1992; Guilherme, 1996), permitindo desta forma obter uma descrição \bar{l} a partir da análise das descrições l_k das cenas c_k .

4.2.1 Editor de Cenários (EdCen)

O Editor de Cenários (**EdCen**) é um editor para criar desenhos ou cenários cujos componentes podem ser tanto objetos básicos, tais como quadrados, retângulos, círculos, etc, quanto elementos pictóricos pré-definidos sobre um tema (Figura 4.4). É um software desenvolvido dentro da noção de programação por objetos (componentes básicos e elementos pictóricos).

O **EdCen** possui vários controles, ferramentas e funções, distribuídos por entre as Barras de Controle (Figura 4.4 no alto), Barra de Ferramentas (Figura 4.4 à esquerda), Barra de Elementos Pictóricos (Figura 4.4 à direita) e Palheta de Cores (Figura 4.4 abaixo da Barra de Controle).

Por ser um software bastante versátil, permite ser utilizado como:

1. Um editor comum de desenhos a partir de suas ferramentas básicas;
2. Um editor de cenários sobre um tema definido, com a utilização de:
 - **Elementos pictóricos**, ou seja, desenhos pré-definidos com componentes de possíveis cenários sobre o tema;
 - **Figuras de fundo** que caracterizem esses possíveis cenários.

Como um editor de cenários, o **EdCen** possui dois módulos distintos:

1. **Montagem:** para definir a figura de fundo e elementos pictóricos a partir de uma base de desenhos pré-definidos, que podem ser utilizadas para exploração da criação de cenários sobre o tema escolhido;
2. **Execução:** para criação das cenas imaginadas pelo usuário.



Figura 4.4 - Exemplo de montagem de cena no EdCen.

4.2.1.1 Som para definir tema

Foi também implementado no **EdCen**, a possibilidade de tocar um som que possa definir o tema a ser utilizado para composição de uma Cena C_k .

A associação de um som ao desenho nos permite investigar a relação da codificação verbal (som), com a visual (montagem de um cenário). Este som pode conter, por exemplo, uma instrução para a montagem de uma cena e é tocado quando o **EdCen** for iniciado, ou repetido sempre que o usuário

desejar. Para isso basta clicar no botão Repete Som na barra de Controle (Figura 4.4). O som inicial praticamente não foi repetido durante a execução das atividades, apesar deste recurso ter sido disponibilizado aos voluntários.

4.2.1.2 Edição de texto

Ao final da composição do cenário, o usuário poderá utilizar um editor de texto simplificado para criar uma história ou desenvolver uma descrição I_k de uma cena c_k criada. Esse editor é ativado a partir do botão “Descrever o Cenário” na Barra de Controle (Figura 4.4). O **EdCen** abrirá uma tela como mostrada na Figura 4.5, para que o indivíduo digite o texto que desejar.

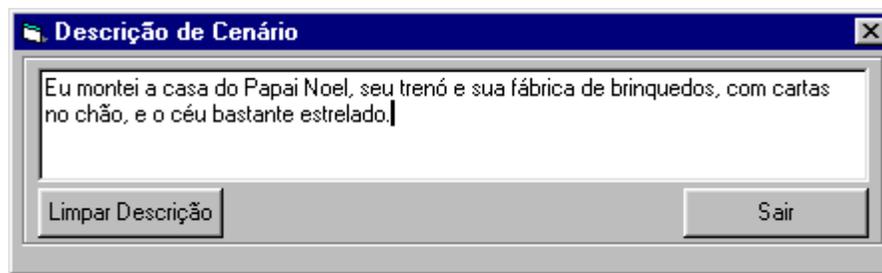


Figura 4.5 - Descritor de Cenário.

4.2.1.3 Bases de dados associadas

Duas bases de dados são utilizadas pelo **EdCen**:

1. **Base de Cenários:** Uma base de dados relacional é usada para armazenar os dados referentes à figura de fundo, som temático e elementos pictóricos que definem cada uma das cenas que serão estudadas.
2. **Base de Usuários:** Uma base de dados relacional que armazena os dados das instâncias dos usuários, isto é, além da figura de fundo e som temático, também os objetos usados pelos usuários para a construção da cena. Em relação aos objetos, são armazenados o nome e informações sobre o tipo (elementos pictóricos, formas geométricas, etc), posicionamento, cor, conteúdo no caso do objeto texto, etc. Esta base de dados também permite organizar usuários em grupos de acordo com o protocolo experimental a ser desenvolvido.

4.2.2 Visualizador e Manipulador de Cenas (VMC)

O Visualizador e Manipulador de Cenas, também chamado de **VMC** (Figura 4.6) é uma ferramenta que foi desenvolvida para visualização, classificação e análise dos cenários criados pelos usuários. Para isso disponibiliza os recursos:

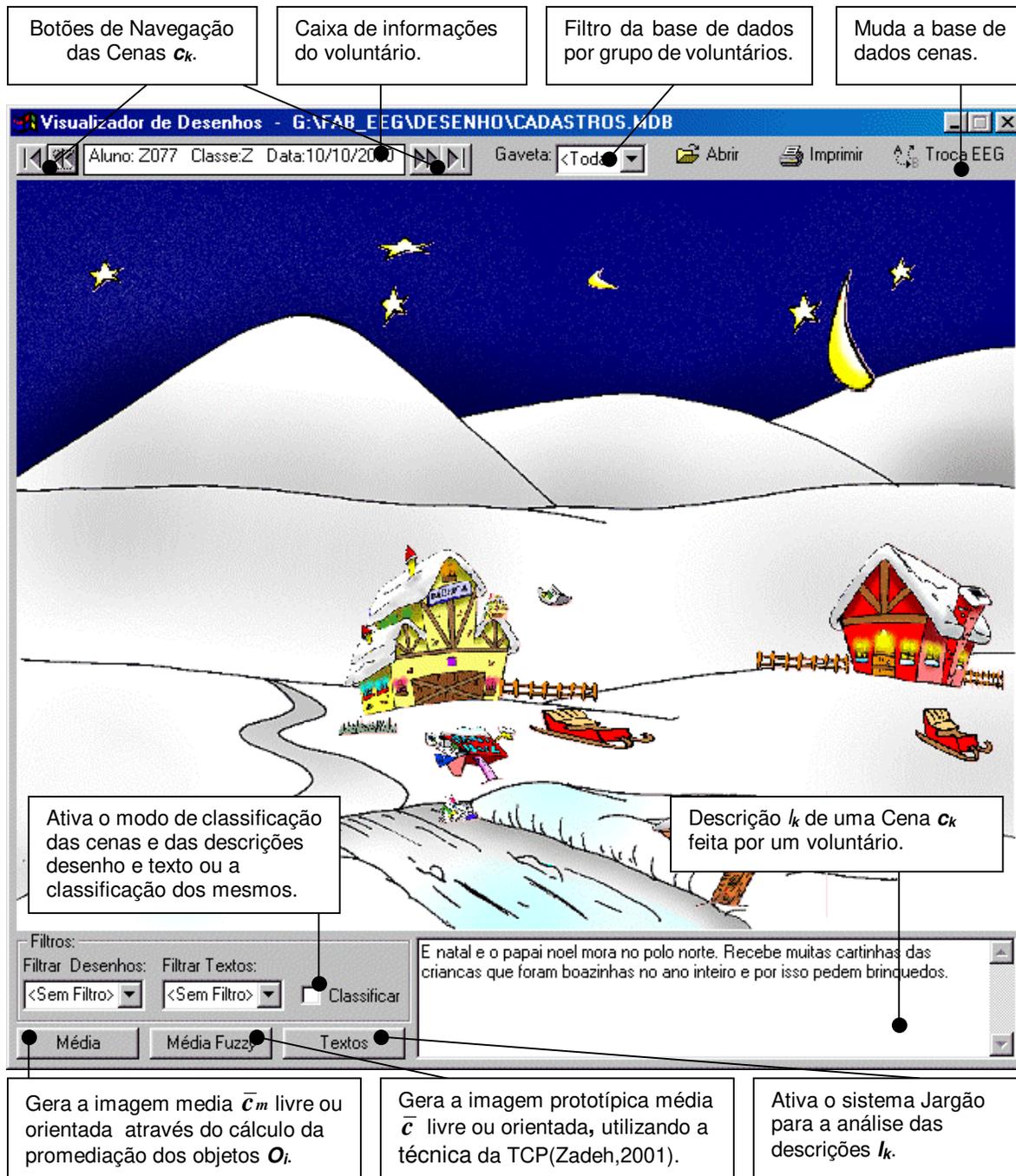


Figura 4.6 - Visualizador e Manipulador de Cenas.

- Filtros – As cenas c_k montadas pelos voluntários podem ser classificadas como boas, médias ou ruins. Após a classificação é possível filtrar a base de dados para que sejam mostrados e entrem no cálculo da imagem prototípica média livre \bar{c}_L ou orientada \bar{c}_O , somente as cenas classificadas como boas, por exemplo. Da mesma forma, é possível classificar e filtrar descrições I_k das cenas. Se necessário, poderá ser aplicado o filtro de qualidade de cenas e qualidade de descrições simultaneamente. Os filtros podem refinar a análise dos cenários, porém neste trabalho, não foi utilizado nenhum filtro de qualidade.
- Geração de imagens prototípicas médias livres \bar{c}_L ou orientadas \bar{c}_O para cada tema, a partir do conjunto C_L ou C_O formado pelas cenas semelhantes c_k , geradas pelo grupo de voluntários. Esta geração é feita através da técnica da TCP, proposta por Zadeh (Zadeh, 2001), ativada através do botão “Média Fuzzy”, mostrado na parte inferior da Figura 4.6, ao lado do botão “Média”.
- Geração de imagens médias livres \bar{c}_{mL} ou orientadas \bar{c}_{mO} para cada tema, a partir do conjunto C_L ou C_O formado pelas cenas semelhantes c_k , geradas pelo grupo de voluntários. Esta geração é feita através do cálculo da promediação das posições e tamanhos dos elementos pictóricos O_i semelhantes do conjunto C , ativada através do botão “Média”, mostrado na parte inferior esquerda da Figura 4.6.
- Análise das descrições I_k das cenas c_k . Esta análise é feita a partir do sistema Jargão (Rocha, 1992; Guilherme, 1996) que interage com o **VMC**. Através desta análise, é possível encontrar trechos de textos s_j comuns entre as diversas descrições I_k e estabelecer relações entre trechos s_j das descrições e o uso de determinados elementos pictóricos.

4.3 Resumo do Capítulo

Neste capítulo foi apresentado o grupo experimental, as atividades desenvolvidas por este grupo e os dois softwares criados:

- um editor de cenários, responsável pela aquisição de cenas e suas respectivas descrições, denominado **EdCen**;
- um visualizador e manipulador das cenas criadas, responsável pelo processamento dos dados obtidos e geração de cenas prototípicas médias, denominado **VMC**.

No capítulo a seguir, serão mostrados os algoritmos para a geração de cenas médias \bar{c}_m e Cenas Prototípicas Médias \bar{c} .

5 Geração de cenas prototípicas médias

Um dos principais objetivos deste trabalho é a geração de cenas prototípicas médias livres \bar{c}_L ou orientadas \bar{c}_O , a partir de conjuntos de cenas livres C_L ou orientadas C_O , formado por cenas c_k semelhantes, desenvolvidas por k voluntários. Espera-se que as cenas \bar{c}_L ou \bar{c}_O geradas, sejam cenas coerentes que representem um consenso da idéia dos voluntários sobre o tema das cenas c_k criadas.

A fim de facilitar a notação, doravante será usado um conjunto C genérico que pode assumir os seguintes valores:

$$(XLIV) \quad C = \{C_L \text{ xor } C_O\},$$

sendo que C pode ser o conteúdo de C_L ou C_O , nunca aparecendo juntos ou sem a presença de um deles, onde C_L corresponde ao conjunto de cenas c_k livres e C_O o conjunto de cenas c_k orientadas.

A partir dessa premissa, podemos dizer que:

$$(XLV) \quad \bar{c}, \text{ é a cena prototípica média do conjunto } C \text{ de cenas semelhantes.}$$

A cena prototípica média \bar{c} é gerada por um algoritmo que tem por base a teoria da TCP (Zadeh, 2001), onde propomos a criação de variáveis lingüísticas de direção e distância, que serão calculadas para cada objeto O_i presente no conjunto C , de cenas semelhantes c_k , em relação a uma localização espacial Z . Para a criação da cena \bar{c} , cada objeto O_i será inserido na direção e distância, em relação a referência Z , que tiveram ocorrência mais freqüente. Este algoritmo será chamado de Promediação Fuzzy.

Para analisar a qualidade da cena \bar{c} gerada, criou-se também um algoritmo que gera uma cena média \bar{c}_m a partir do mesmo conjunto de cenas C , através de um cálculo de promediação convencional que coloca cada elemento pictórico O_i na posição média de todas as ocorrências deste objeto O_i no conjunto C .

Serão mostrados no capítulo 7 - Resultados, as cenas prototípicas médias \bar{c} e as cenas médias \bar{c}_m obtidas.

5.1 Algoritmo de promediação convencional para geração da cena média \bar{c}_m

O objetivo deste algoritmo é gerar uma cena média \bar{c}_m a partir de um conjunto \mathbf{C} de cenas \mathbf{c}_k semelhantes, que será composta por todos os objetos \mathbf{O}_i usados e colocados em suas coordenadas médias.

O algoritmo de promediação é ativado a partir do botão “Média” do **VMC** (Figura 4.6) e quando ativado é mostrado o formulário da Figura 5.1. Neste formulário, define-se qual conjunto \mathbf{C} de cenas semelhantes \mathbf{c}_k deverá ser processado (Abertura, Cena Mais Engraçada, Fábrica de Brinquedos) e quais os limiares de corte de frequência que se deseja aplicar.

Os limiares de corte definem a quantidade mínima e a máxima de objetos que devam entrar na média. Esta quantidade é adquirida pelo número de indivíduos k que utilizaram determinado elemento pictórico, e não o número de elementos pictóricos \mathbf{O}_i usados no geral, já que um indivíduo pode utilizar o mesmo elemento pictórico \mathbf{O}_i mais de uma vez.

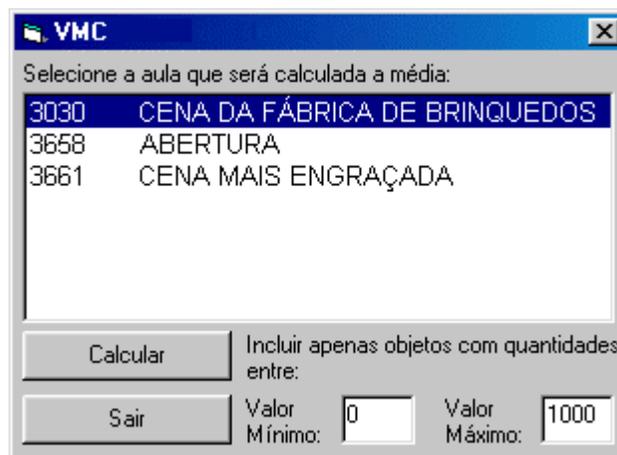


Figura 5.1 - Selecionador de conjunto de cenas e configurador de limiares de corte.

A cena prototípica média \bar{c}_m é, portanto, obtida da seguinte forma:

- Calcula-se, para cada cena \mathbf{c}_k , a média das posições do mesmo elemento pictórico \mathbf{O}_i que porventura tenha sido utilizado mais de uma vez, obtendo assim uma imagem \mathbf{c}_k' . Por exemplo, se um usuário k utilizou na cena \mathbf{c}_k vários elementos pictóricos \mathbf{O}_i “estrela”, é calculada para esta cena \mathbf{c}_k a posição média do objeto “estrela”. Após o término desta etapa, cada objeto \mathbf{O}_i terá apenas uma ocorrência na cena \mathbf{c}_k' e para cada cena \mathbf{c}_k será gerado uma cena \mathbf{c}_k' ;

- A partir das cenas médias c_k' geradas, calcula-se as coordenadas médias de posição para cada objeto O_i utilizado no conjunto de cenas C' que é o conjunto de cenas c_k' geradas;
- Calcula-se a frequência de uso dos objetos O_i no conjunto C' ;
- Elimina-se da lista os objetos O_i que tenham uma frequência inferior e superior ao limite de corte definido para análise;
- Compõe-se a cena média \bar{c}_m inserindo cada objeto O_i em sua coordenada média.

A seguir, a Figura 5.2 mostra o fluxograma (Pressman, 1995) do algoritmo da geração da cena média \bar{c}_m .

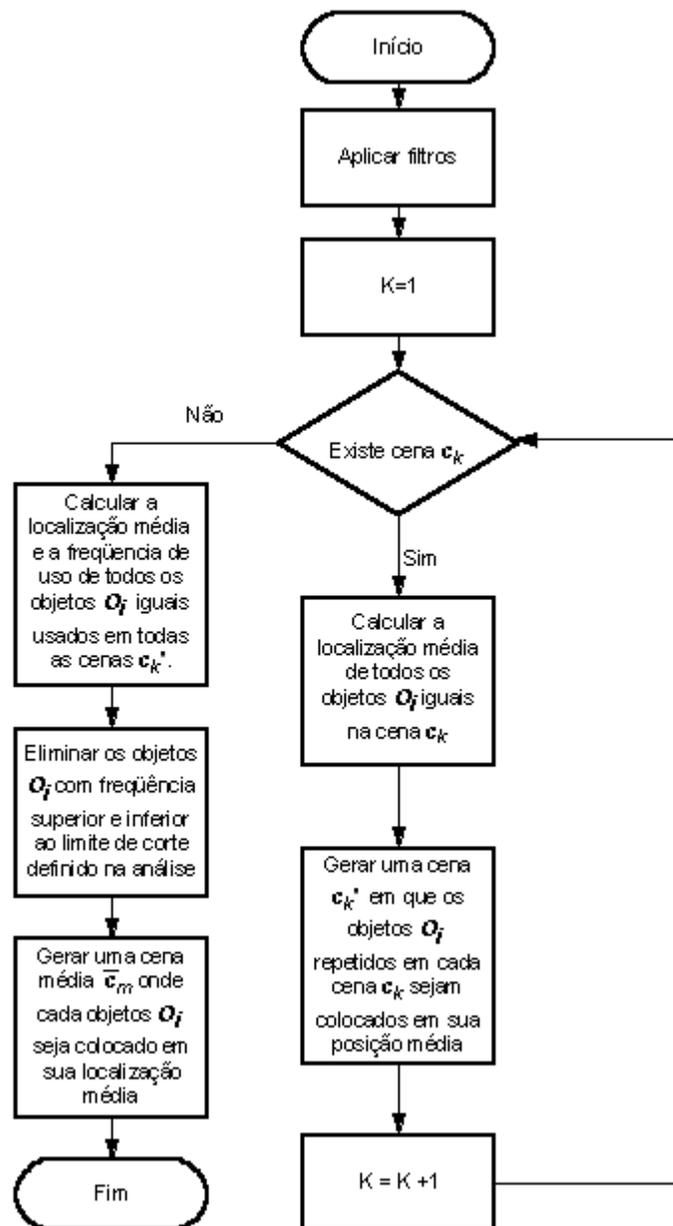


Figura 5.2 - Fluxograma do algoritmo da geração da cena média \bar{c}_m .

5.2 Algoritmo de Promediação Fuzzy para geração da cena prototípica média \bar{c} , utilizando a proposta da TCP

Como discutido anteriormente, a análise de uma cena c_k é realizada a partir do cálculo da localização do conjunto de objetos O_i em relação a uma referência Z . Isso pode ser formalizado como na equação (IV):

$$c_k = [q_1, q_2, \dots, q_n], \text{ onde:} \\ q_i = (p_1, p_2) = \text{Localização}(O_i, Z)$$

o que significa que a questão q é repetida para todos os objetos presentes na cena c_k e Z pode ser um elemento pictórico específico O_r ou qualquer ponto no espaço XY .

Dado um conjunto $C = \{c_1, c_2, \dots, c_m\}$ de cenas semelhantes, é possível calcular uma cena prototípica \bar{c} composta pelos objetos O_i mais freqüentes em C .

Desse modo, \bar{c} será caracterizada pela disposição dos objetos O_i em relação a Z , porém ordenada por sua freqüência em C , isto é, as localizações mais freqüentes terão prioridade no seu posicionamento em \bar{c} .

A freqüência de posicionamento de objetos pode ser obtida a partir dos histogramas dos possíveis valores assumidos por cada termo lingüístico de distância (Muito Perto, Perto, Média, Longe, Muito Longe) e direção (Norte, Sul, Leste, Oeste, Noroeste,...). Pode-se supor que as proposições que obtiveram maior freqüência em cada histograma correspondem a um consenso sobre a localização dos objetos O_i em C , e por isso devem definir \bar{c} .

Para ativar o cálculo da cena prototípica média \bar{c} , deve-se clicar no botão "Média Fuzzy" do **VMC** (Figura 4.6) que exibirá o formulário da Figura 5.1. Neste formulário, define-se qual conjunto C de cenas semelhantes c_k deverá ser processado (Abertura, Cena Mais Engraçada, Fabrica de Brinquedos) e quais os limiares de corte de freqüência que se deseja aplicar.

Os limiares de corte, assim como explicado anteriormente, definem a quantidade mínima e a máxima de objetos que devem entrar na média. Esta quantidade é adquirida pelo número de indivíduos k

que utilizaram determinado elemento pictórico, e não o número de elementos pictóricos O_i usados no geral, já que um indivíduo pode utilizar o mesmo elemento pictórico mais de uma vez.

Uma vez definido isso, o **VMC** (Figura 4.6) requisitará que seja informada a referência Z . Tendo sido definida a referência Z , o **VMC** então calculará para cada objeto O_i , a direção e a distância em relação à referência Z .

De posse destes dados será possível então calcular a imagem prototípica média \bar{c} .

5.2.1 Obtenção da referência Z

Para fixar uma referência Z , o **VMC** exibe a tela mostrada na Figura 5.3.

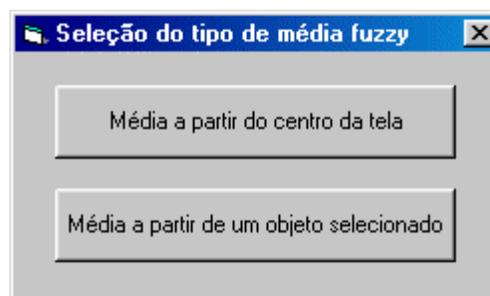


Figura 5.3 - Localização Z .

Como proposto anteriormente, a referência Z pode ser um elemento pictórico específico O_r ou qualquer ponto no espaço XY .

Se for escolhido o botão "Média a partir do centro da tela" (Figura 5.3), então a referência Z será o centro da tela, caso contrário será exibida a tela da Figura 5.4.

A tela mostrada na Figura 5.4, exibe todos os elementos pictóricos O_i disponíveis, além do número de voluntários k que usaram cada objeto O_i , sendo que esta informação é exibida na caixa de texto logo abaixo da imagem dos elementos pictóricos O_i .

Ao clicar num objeto O_i , a cena prototípica média \bar{c} será montada a partir da posição central da média deste objeto O_i no conjunto de cenas c_k semelhantes, ou seja, a cena prototípica média \bar{c} será

montada a partir do ponto de vista deste objeto O_i . Para a obtenção da referência Z , devem ser realizados os seguintes procedimentos:

- Dado o elemento pictórico O_i selecionado, calcula-se para cada cena c_k , a média das posições dos elementos pictóricos O_i iguais, que porventura tenham sido utilizado mais de uma vez. Desta forma para cada cena c_k existirá um ponto P_k , que será a posição média do objeto O_i selecionado;
- Para o conjunto C de cenas c_k semelhantes, calcula-se a média dos pontos P_k , gerando assim uma posição média P . A referência Z será a posição P .

Pode-se também retirar, por qualquer razão que seja, algum objeto O_i do cálculo da cena prototípica média \bar{c} , marcando a caixa de checagem “Retira Objeto” (Figura 5.4).



Figura 5.4 - Tela de seleção do elemento pictórico O_i que será tomado como referência Z para a geração da cena prototípica média \bar{c} .

5.2.2 Distância de um objeto O_i em relação a uma referência Z

O valor numérico da distância euclidiana $d_{i,j}$ entre um objeto O_i e uma referência Z é referenciado a cada um dos cinco termos lingüísticos utilizados. Esse processo associa um valor de pertinência dessa distância a cada um dos termos lingüísticos, de acordo com as funções de pertinência definidas para cada um desses termos. Essa distância referenciada $d_{i,j}$ assumirá como valor o termo lingüístico de maior pertinência. Lembrando que, $d_{i,j}$ (em negrito e itálico), representa uma distância através de um termo lingüístico, e $d_{i,j}$ (somente em negrito), representa uma distância numérica (crisp). Portanto:

(XLVI)

$$d_{i,j} = \arg(\mathbf{v} \mu_k(d_{i,j})), \text{ onde,}$$

$$L = \{\text{Muito Perto, Perto, Médio, Longe, Muito Longe}\} \text{ e } k \in L$$

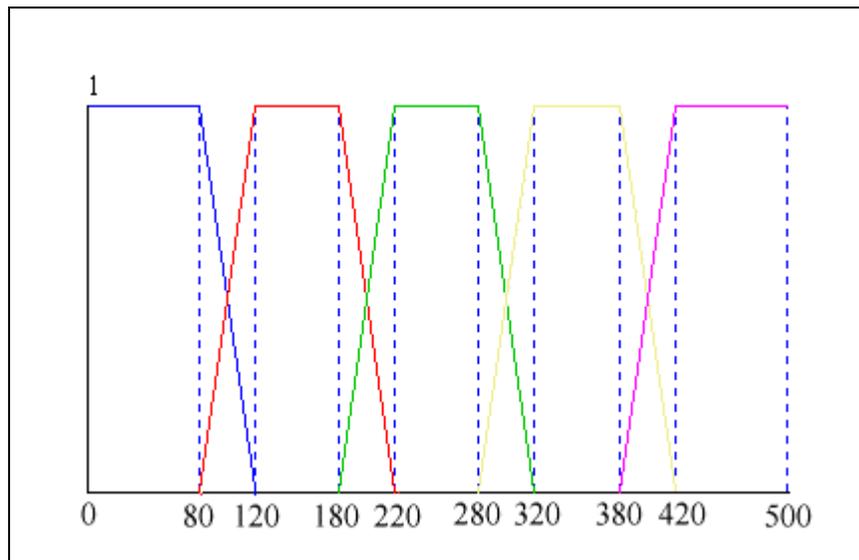


Figura 5.5- Funções de pertinência para a análise da distância entre um objeto O_i e a referência Z .

Além do termo lingüístico de maior pertinência, é guardado também numa base de dados o grau de pertinência de cada termo lingüístico da variável lingüística de distância. Essa informação será usada posteriormente na defuzzificação.

O gráfico da Figura 5.5, mostra as funções de pertinência usadas para classificação da distância euclidiana $d_{i,j}$ dos objetos O_i , usados no conjunto C de cenas c_k , sendo que cada uma das retas correspondem, respectivamente, aos termos lingüísticos: Muito Perto, Perto, Médio, Longe, Muito Longe.

A Figura 5.6 mostra um exemplo de classificação de distância. No exemplo, a referência Z é o centro do elemento pictórico O_0 "Trenó". O valor numérico da distância euclidiana $d_{i,j}$, entre o elemento

pictórico O_1 "Casa" e a referência Z é 325, sendo que a tela de fundo possui a escala 0..1000 x 0..1000. Portanto, aplicando as funções de pertinência da Figura 5.5 à distância euclidiana $d_{ij} = 325$, obteremos como resultado o termo lingüístico de maior pertinência, no caso $d_{ij} = \text{Longe}$.

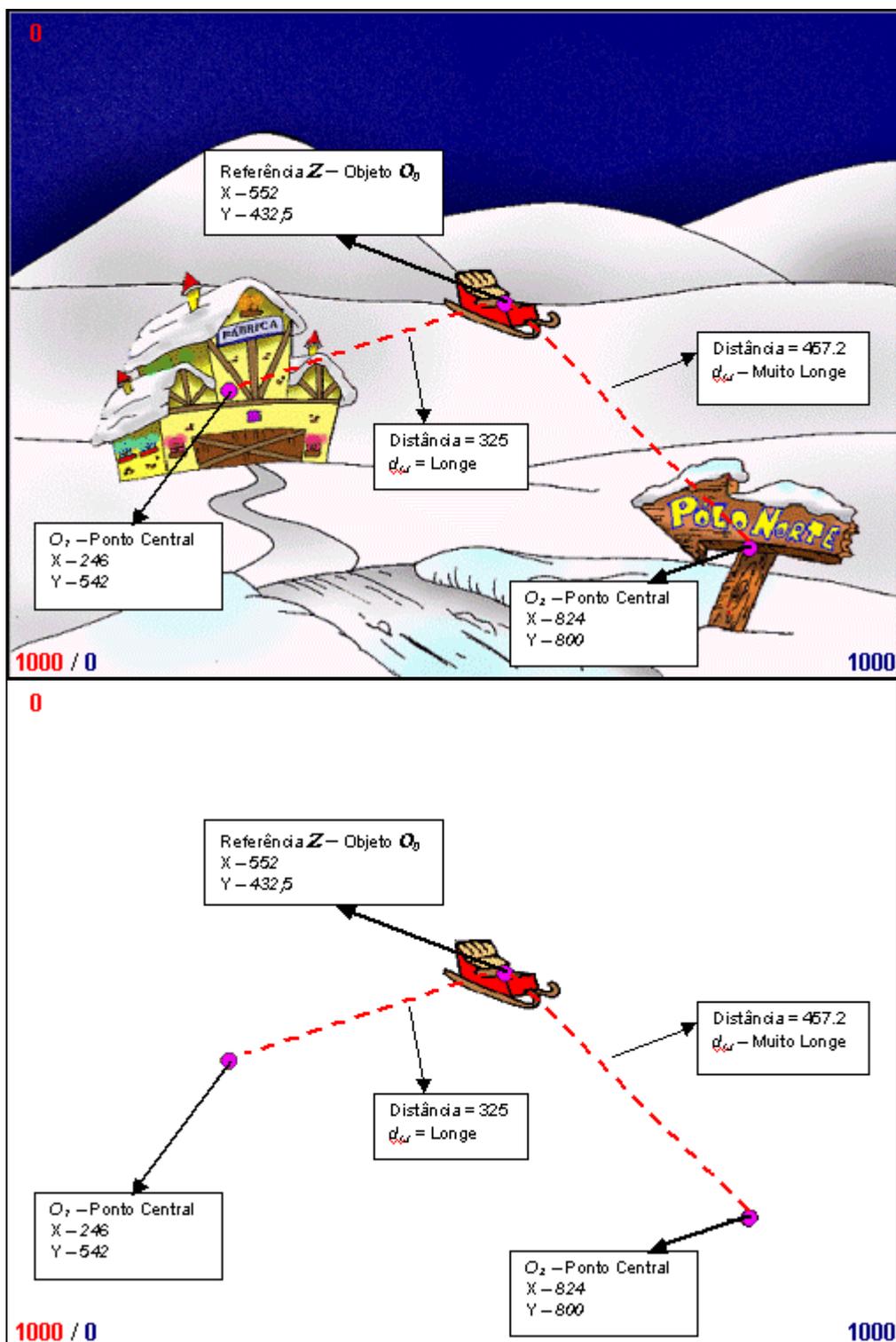


Figura 5.6 - Exemplo da classificação de elementos pictóricos em relação à distância.

A Tabela 5.1 mostra o valor de pertinência para cada variável lingüística adquirida a partir da distância euclidiana $d_{i,j}$ entre a referência Z e os objetos O_1 e O_2 da Figura 5.6 .

Valor de pertinência de cada termo lingüístico a partir da distância euclidiana entre o objeto O_1 e a referência Z .		Valor de pertinência de cada termo lingüístico a partir da distância euclidiana entre o objeto O_2 e a referência Z .	
$\mu V_{\text{MuitoPerto}}(325)$	0	$\mu V_{\text{MuitoPerto}}(457,2)$	0
$\mu V_{\text{Perto}}(325)$	0	$\mu V_{\text{Perto}}(457,2)$	0
$\mu V_{\text{Médio}}(325)$	0	$\mu V_{\text{Médio}}(457,2)$	0
$\mu V_{\text{Longe}}(325)$	1,0	$\mu V_{\text{Longe}}(457,2)$	0
$\mu V_{\text{MuitoLonge}}(325)$	0	$\mu V_{\text{MuitoLonge}}(457,2)$	1,0

Tabela 5.1 - Valor de pertinência de cada termo lingüístico a partir da distância euclidiana $d_{i,j}$ entre um objeto O_i e a referência Z .

É possível, desta forma, criar um histograma de distância que mostre o número de ocorrências de cada uma das cinco variáveis lingüísticas (Muito Perto, Perto, Médio, Longe, Muito Longe) para um determinado objeto O_i , num conjunto C de cenas c_k semelhantes, como mostra a Figura 5.7.

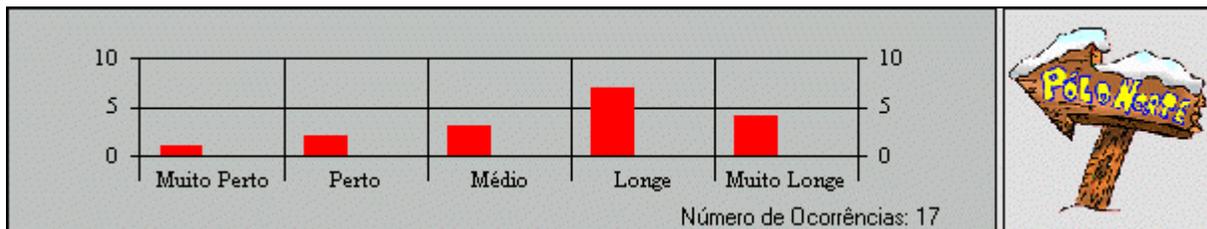


Figura 5.7 - Exemplo de um histograma de distância de um objeto O_i .

5.2.2.1 Histograma de distância

O **VMC** possui também um visualizador dos histogramas de distâncias (Figura 5.8). Este visualizador mostra todos os histogramas das variáveis lingüísticas de distância, geradas para cada elemento pictórico O_i .

Mesmo que a referência Z seja definida como sendo a posição central média de um elemento pictórico O_r , este objeto O_r também fará parte do histograma de distância, sendo mostrado neste histograma o número de ocorrências das variáveis lingüísticas de cada objeto O_r em todas as cenas c_k em relação à referência Z , que é sua posição média. Este histograma mostra a variância da posição de O_r .

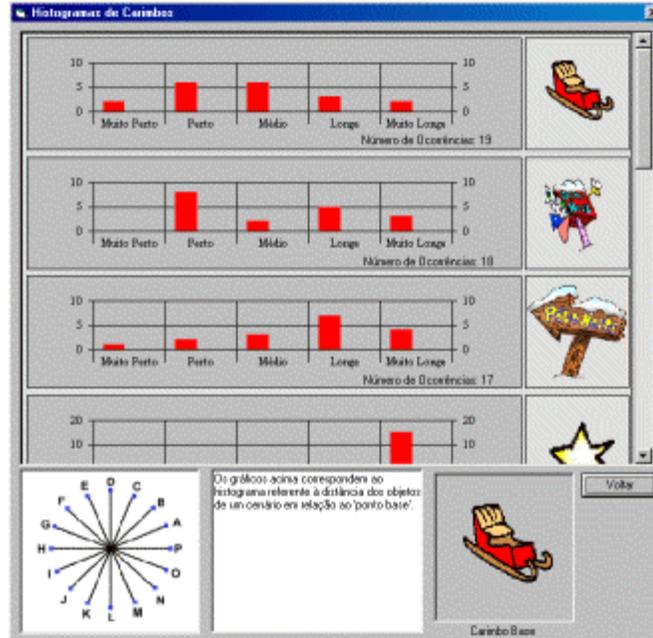


Figura 5.8 - Gráfico dos histogramas de distâncias.

5.2.2.2 Editor de Retas Fuzzy

O VMC também disponibiliza um Editor de Retas Fuzzy (Figura 5.9), para que se possa personalizar os valores de cada função de pertinência, para cada conjunto C de cenas c_k . Através da personalização das retas fuzzy para cada conjunto C de cenas c_k , é possível tentar obter a melhor cena prototípica média \bar{c} .

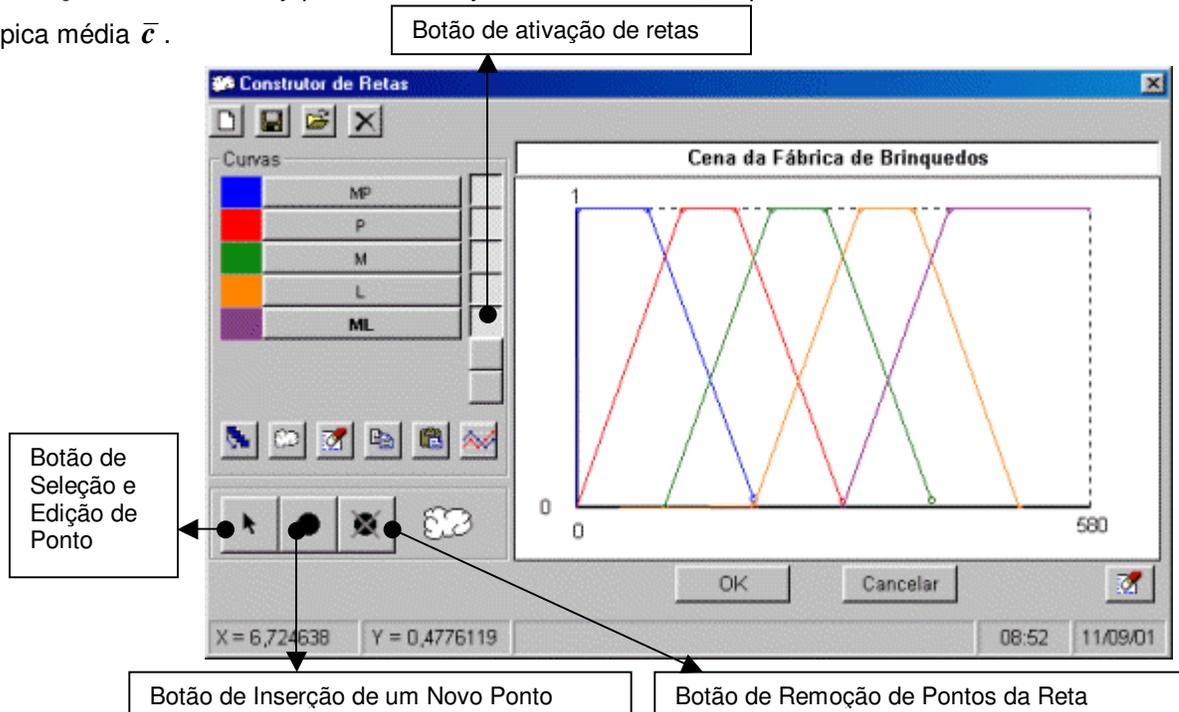
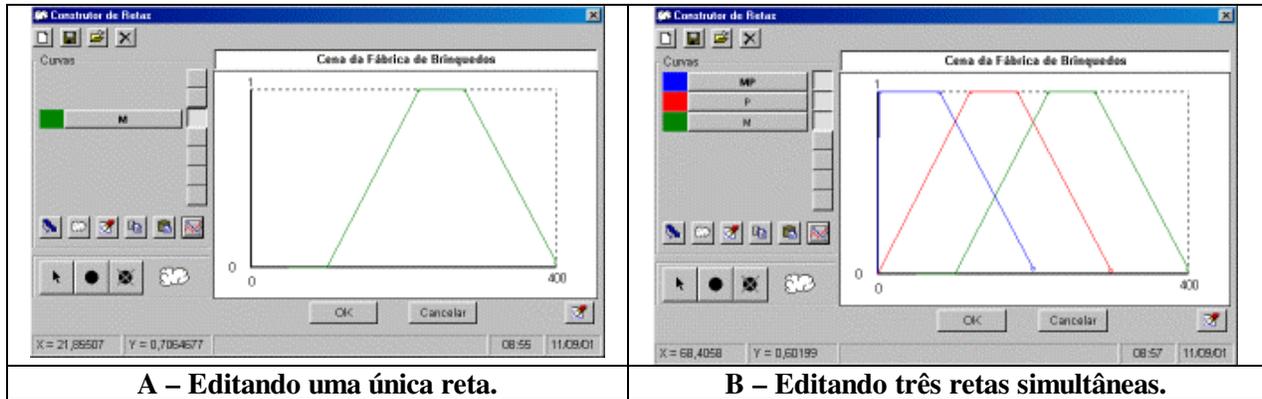


Figura 5.9 - Formulário de montagem das retas fuzzy.



A – Editando uma única reta.

B – Editando três retas simultâneas.

Figura 5.10 - Editando Retas Fuzzy.

O Editor de Retas Fuzzy permite editar as retas de forma individual ou em conjunto, conforme mostram as Figura 5.10 – A e Figura 5.10 - B, ativando-as ou desativando-as a partir do botão de ativação de retas (Figura 5.9).

5.2.3 Direção de um objeto O_i em relação a uma referência Z

Para o cálculo da direção entre o centro de um objeto O_i e a referência Z , usaremos a mesma tecnologia utilizada pelo sistema **Sensor** para detecção de bordas de objetos, ou seja, as equações aqui utilizadas são as mesmas discutidas no item 3.4.3 - Um Sistema Visual Formal.

Desta forma, a localização l_{ij} entre um objeto O_i e uma referência Z , pode ser definida como na equação (XXXIII):

$$l_{ij} = \{ l_x, l_y \}$$

$$l_x = x_i - x_j, l_y = y_i - y_j$$

Utilizando o dicionário D_3 mostrado em (XXXIV):

$$D_3 = \{ p = \text{positivo}, n = \text{negativo}, z = \text{zero}, a = \text{vazio}, b = \text{pequeno}, c = \text{médio}, d = \text{grande}, e = \text{enorme} \}$$

e as relações de localização mostrada na equação (XXXV):

Negativo	l_x ou $l_y < 0$
Positivo	l_x ou $l_y > 0$
Zero	l_x ou $l_y = 0$
Vazio	$l_x / l_y < \alpha_1$
Pequeno	$\alpha_1 \leq l_x / l_y < \alpha_2$
Médio	$\alpha_2 \leq l_x / l_y < \alpha_3$
Grande	$\alpha_3 \leq l_x / l_y < \alpha_4$
Enorme	$l_x / l_y \geq \alpha_4$

é possível gerar uma sentença s_1 que descreva a localização entre o objeto O_i e a referência Z . Onde s_1 é uma sentença suportada por uma gramática G_1 , isto é:

$$(XLVII) \quad s_1 = \{ c_{x1} c_{y1} c_1 \}^j,$$

tal que:

$$(XLVIII) \quad c_{x1}, c_{y1} \in \{ p, n, z \}, c_1 \in \{ b, c, d, e \},$$

em outras palavras, s_1 descreve as direções entre dois pontos através de uma gramática em que c_1 classifica um ângulo e $c_{x1} c_{y1}$ identifica o seu quadrante.

A sentença s_1 pode ser reescrita em uma outra sentença s_2 suportada pela gramática G_2 , utilizando o mesmo dicionário usado na equação (XXXVIII):

$$D_4 = \{ A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P \},$$

onde os valores presentes no dicionário D_4 são obtidos através da equação (XXXIX):

$$\left\{ \begin{array}{cccc} p^*a \rightarrow P & ppb \rightarrow A & ppc \rightarrow B & ppd \rightarrow C \\ *pe \rightarrow D & npd \rightarrow E & npc \rightarrow F & npb \rightarrow G \\ n^*a \rightarrow H & nnb \rightarrow I & nnc \rightarrow J & nnd \rightarrow K \\ *ne \rightarrow L & pnd \rightarrow M & pnc \rightarrow N & pnb \rightarrow O \end{array} \right\},$$

sendo que “*”, significa “qualquer direção”. A Figura 5.11, mostra a representação gráfica do dicionário D_4 , de acordo com as regras exibidas na equação (XXXIX).

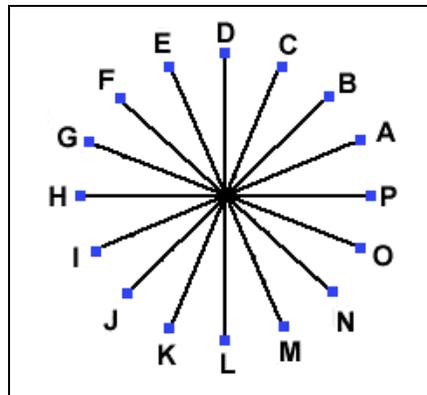


Figura 5.11 - Representação gráfica das direções do dicionário D_4 .

A Figura 5.12 mostra um exemplo onde a referência Z é o centro do elemento pictórico O_0 “Trenó”. Este exemplo possui também dois outros elementos pictóricos, O_1 e O_2 , que terão sua classificação de direção feita da seguinte forma:

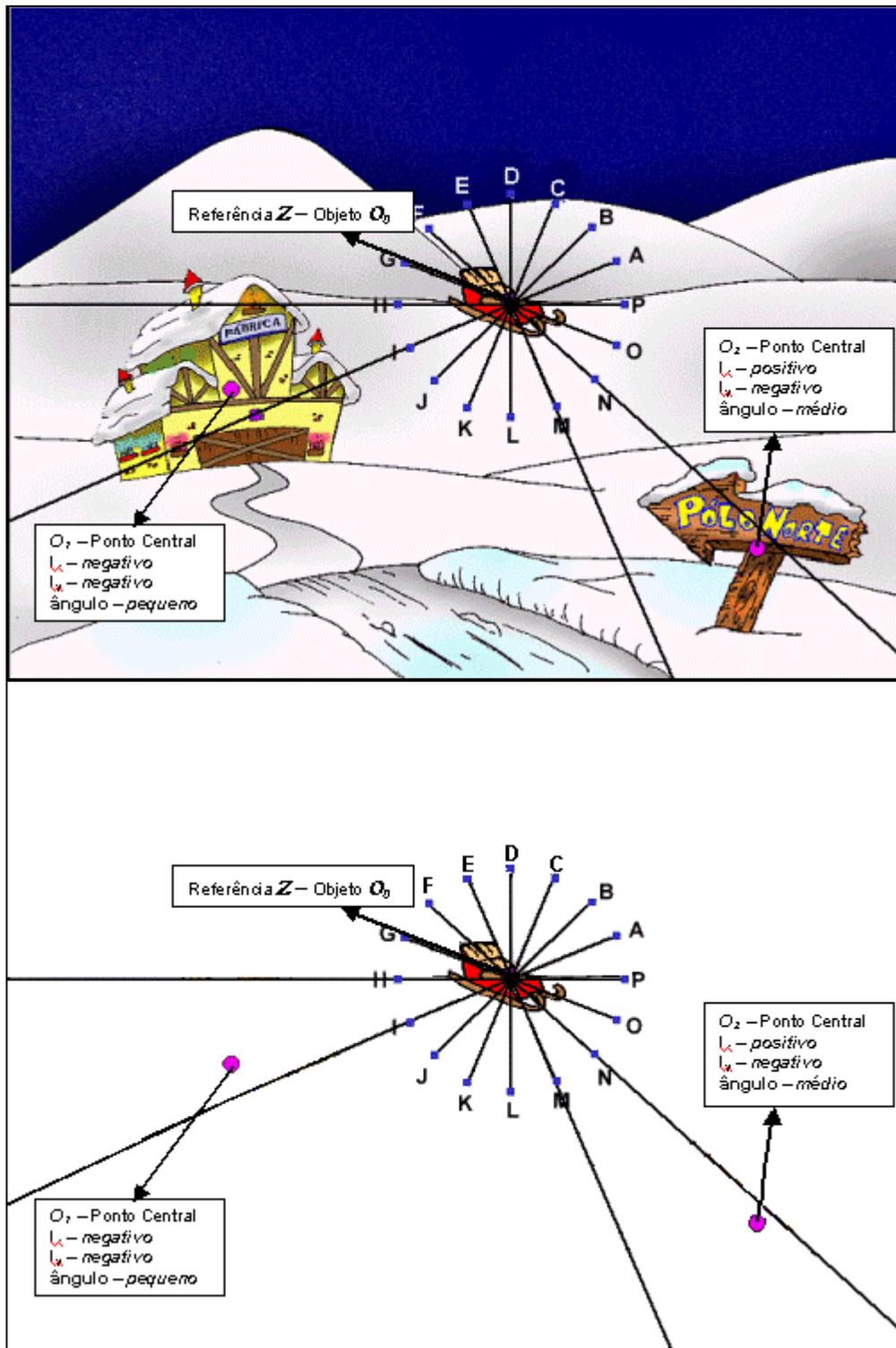


Figura 5.12 - Exemplo da classificação da direção de um elemento pictórico O_i em relação a referência Z .

- O elemento pictórico O_1 (casa), terá sua direção marcada como “I”, já que $I_x < 0$, $I_y < 0$ e o ângulo entre a referência Z e o objeto O_1 é pequeno, tendo a sentença s_1 escrita como “nmb”.
- O elemento pictórico O_2 (placa), terá sua direção marcada como “N”, já que $I_x > 0$, $I_y < 0$ e o ângulo entre a referência Z e o objeto O_2 é médio, tendo a sentença s_1 escrita como “pnc”.

Portanto, é possível criar um histograma de direções que mostre o número de ocorrências de cada um dos dezesseis termos lingüísticos do dicionário D_4 para um determinado objeto O_i , num conjunto C de cenas c_k semelhantes, como mostra a Figura 5.13.

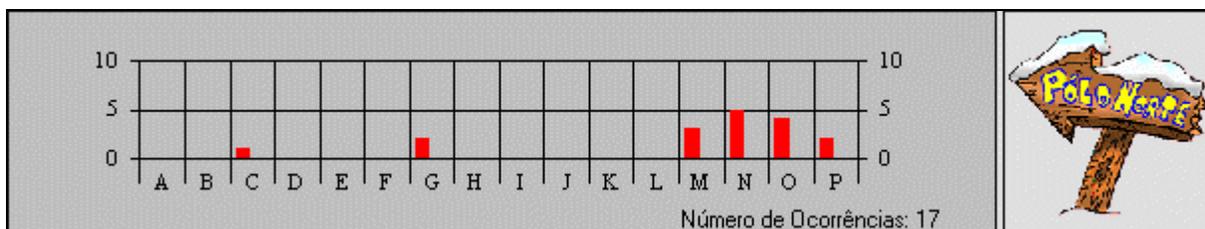


Figura 5.13 - Exemplo de um histograma de direção de um objeto O_i .

5.2.3.1 Histograma de direção

O **VMC** possui também um visualizador dos histogramas de direções (Figura 5.14). Este visualizador mostra todos os histogramas dos termos lingüísticos do dicionário D_4 de direções, geradas para cada elemento pictórico O_i .

Mesmo que a referência Z seja definida como sendo a posição central de um elemento pictórico O_r , este objeto O_r também fará parte do histograma de direção, sendo mostrado neste histograma o número de ocorrências das variáveis lingüísticas de cada objeto O_r em todas as cenas c_k , em relação a referência Z , que é sua posição média. Esse histograma mostra a variância da posição de O_r .

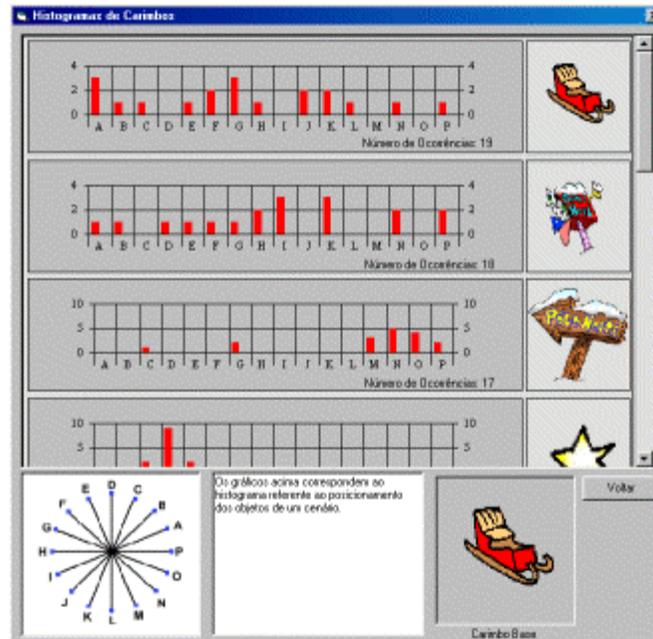


Figura 5.14 - Gráfico dos histogramas de direções.

5.2.4 Geração da cena prototípica média \bar{c}

Para a geração da cena prototípica média \bar{c} , segue-se o seguinte protocolo:

- Aplicam-se os filtros desejados, seleciona-se para qual conjunto \mathbf{C} (Abertura, Cena Mais Engraçada, Fábrica de Brinquedos) de cenas \mathbf{c}_k semelhantes deseja-se calcular a cena prototípica média \bar{c} e fixa-se os limiares de corte, que definem a quantidade mínima e a máxima de objetos que devem entrar na média. Esta quantidade é adquirida pelo número de indivíduos k que utilizaram determinado elemento pictórico \mathbf{O}_i e não o número de elementos pictóricos \mathbf{O}_i usados no geral, já que um indivíduo pode utilizar o mesmo objeto \mathbf{O}_i mais de uma vez;
- Define-se a referência \mathbf{Z} , como explicado no item 5.2.1 – Obtenção da referência \mathbf{Z} ;
- Para cada objeto \mathbf{O}_i de todas as cenas semelhantes \mathbf{c}_k , calcula-se a distância $d_{i,j}$ entre o objeto \mathbf{O}_i e a referência \mathbf{Z} , obtendo como resultado um termo lingüístico de distância, como foi explicado no item 5.2.2 – Distância de um objeto \mathbf{O}_i em relação a uma referência \mathbf{Z} ;

- Para cada objeto O_i de todas as cenas semelhantes c_k , calcula-se o termo lingüístico referente à direção entre o objeto O_i e a referência Z , como explicado no item 5.2.3 – Direção de um objeto O_i em relação a uma referência Z ;
- Tendo para cada objeto O_i de todas as cenas c_k semelhantes, os valores dos termos lingüísticos de direção e distância, são calculados os histogramas de direção e distância, somando-se as ocorrências dos termos lingüísticos das variáveis lingüísticas de direção e distância para cada objeto O_i (Figura 5.15).

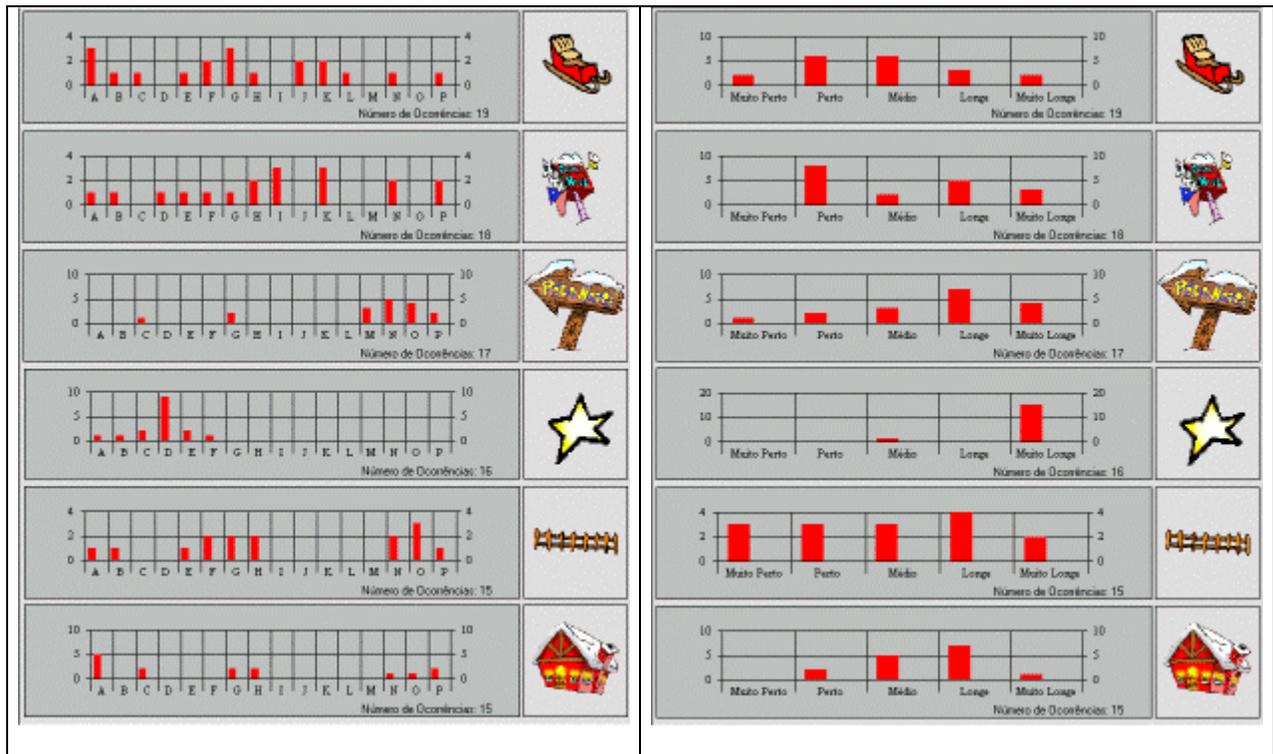


Figura 5.15 - Exemplos de histogramas de direção e distância de objetos O_i .

- A colocação dos objetos O_i na tela de fundo da cena prototípica média \bar{c} , que está sendo montada, dá-se na ordem crescente de quantidade de indivíduos k que utilizaram cada objeto O_i , ou seja, os objetos mais usados são os primeiros a serem inseridos. Caso tenha sido escolhido um objeto O_r para ser a referência Z , este objeto O_r é o primeiro a ser inserido e é colocado na posição da referência Z .

Para a inclusão dos objetos O_i na cena prototípica média \bar{c} , seguem-se os passos:

1. O posicionamento do objeto O_i em \bar{c} é feito a partir de sua distância euclidiana $d_{i,z}$ e direção $l_{i,z}$, onde:

- Inicialmente, através da análise do histograma de distância, seleciona-se a distância lingüística $d_{i,z}$ mais freqüente. A recodificação da distância euclidiana $d_{i,z}$ (defuzzificação) a partir da distância referenciada $d_{i,z}$ é definida da seguinte maneira:

a. Percorre-se todas as cenas c_k e para cada objeto O_i que possua $d_{i,j}$ igual à distância lingüística mais freqüente, é feito o cálculo de defuzzificação, como mostrado no tópico 3.1.2, ou seja, é usado o cálculo da equação (X):

$$d_{i,j} = (q_{i-1} * \mu_{i-1} + q_i * \mu_i + q_{i+1} * \mu_{i+1}) / \mu_{i-1} + \mu_i + \mu_{i+1}.$$

Onde:

- μ_i , é o valor da função de pertinência de maior grau;
 - q_i , é o ponto médio da região de máximo valor da função de pertinência de maior grau (μ_i);
 - i , é o termo lingüístico referente à função de pertinência de maior grau. Os valores que i pode assumir são: “Muito Perto”, “Perto”, “Médio”, “Longe”, “Muito Longe”.
- b. Dessa maneira, para cada objeto O_i que possua $d_{i,j}$ igual à distância lingüística mais freqüente, será obtido uma distância euclidiana $d_{i,j}$. O valor final da distância euclidiana, onde deverá ser inserido o objeto, é a posição média de todos as distâncias $d_{i,j}$ obtidas.

Conseqüentemente, $q_{i-1} * \mu_{i-1}$ diz respeito ao cálculo da função de pertinência referente ao termo lingüístico anterior ao termo lingüístico de maior grau de pertinência, multiplicado pelo ponto médio da região de máximo valor da função de pertinência (μ_{i-1}). Se $i = \text{“Muito Perto”}$, esta etapa é ignorada. Já $q_{i+1} * \mu_{i+1}$ diz respeito ao cálculo da função de pertinência referente ao termo lingüístico posterior ao termo lingüístico de maior grau de pertinência, multiplicado pelo ponto médio da região de máximo valor da função de pertinência (μ_{i+1}). Se $i = \text{“Muito Longe”}$, esta etapa é ignorada.

Portanto, supondo que $d_{i,j} = \text{Médio}$, a defuzzificação seria feita da seguinte forma:

$$d_{i,j} = (q_{\text{Perto}} * \mu_{\text{Perto}} + q_{\text{Médio}} * \mu_{\text{Médio}} + q_{\text{Longe}} * \mu_{\text{Longe}}) / (\mu_{\text{Perto}} + \mu_{\text{Médio}} + \mu_{\text{Longe}})$$

Devido ao formato das funções de pertinência usadas neste trabalho, sempre que o grau de pertinência for igual a um, o resultado da defuzzificação será igual a q_i , já que μ_{i-1} e μ_{i+1} terão o valor zero.

A Figura 5.16 mostra as funções de pertinência para cada um dos cinco termos lingüísticos utilizados: Muito Perto, Perto, Médio, Longe, Muito Longe.

Vamos supor que deseja-se efetuar a defuzzificação do valor lingüístico $d_{i,z} = \text{Médio}$, cujo valor da função de pertinência é $\mu_{\text{Médio}} = 0,8$. A partir desta informação, devemos analisar qual o valor da função de pertinência do termo lingüístico anterior e posterior a *médio*. Desta forma, obteremos os seguintes valores: $\mu_{\text{Perto}} = 0,2$ e $\mu_{\text{Longe}} = 0$.

Portanto, o valor de $d_{i,z}$ referente à defuzzificação do valor referenciado $d_{i,z}$ é:

$$d_{i,j} = (q_{\text{Perto}} * \mu_{\text{Perto}} + q_{\text{Médio}} * \mu_{\text{Médio}} + q_{\text{Longe}} * \mu_{\text{Longe}}) / (\mu_{\text{Perto}} + \mu_{\text{Médio}} + \mu_{\text{Longe}})$$

$$d_{i,j} = (150 * 0,2 + 250 * 0,8 + 350 * 0) / (0,2 + 0,8 + 0) = 230$$

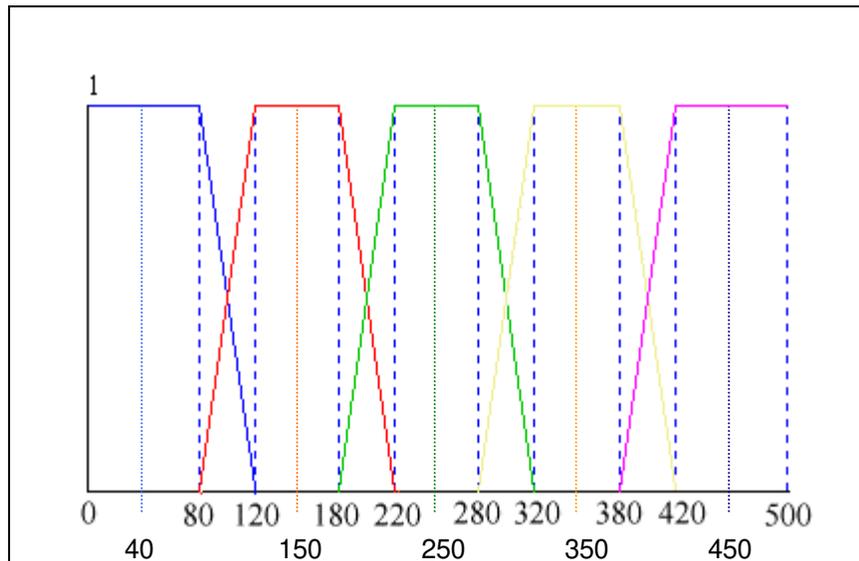


Figura 5.16 - Defuzzyficação das variáveis lingüísticas de distância.

- Depois disso, através da análise do histograma de direção, seleciona-se a direção mais freqüente e a recodificação desta direção $I_{i,z}$ resulta em um valor do ângulo α definido na Figura 5.11 para a variável lingüística correspondente. Na Figura 5.11, o ângulo de cada direção cresce 22,5°, a partir da variável lingüística “A” e é incrementado no sentido anti-horário.

2. Se a posição estiver vazia, ou seja, se não houver nenhum elemento pictórico O_x , já colocado na distância $d_{i,j}$ e no ângulo α , o objeto O_i é inserido nesta direção e distância, e a posição é marcada como ocupada;
3. Se já houver um elemento pictórico nesta posição, fixa-se a distância $d_{i,j}$ mais freqüente e tenta-se inserir o objeto O_i no segundo termo lingüístico de direção de maior freqüência do histograma de direção. Feito isso, se repetem os procedimentos 1 e 2 descritos acima. Caso não se obtenha êxito na inserção do objeto O_i , é repetido o procedimento 3 para o termo lingüístico de direção de terceira ou quarta maior freqüência;
4. Se ainda não for encontrada uma posição disponível, tenta-se então fixar a direção mais freqüente, ou seja, o ângulo α , variando desta vez as distâncias mais freqüentes e repetindo os passos 1, 2 e 3 descritos acima para tentar inserir um objeto O_i na cena prototípica média \bar{c} . Dessa forma, é baixa a probabilidade de não se encontrar uma posição disponível;
5. Se mesmo assim não houver posições disponíveis, então o elemento pictórico não é inserido.

Ao inserir um objeto O_i numa determinada direção e distância, e este ultrapassar os limites da tela, então o elemento pictórico é colocado no limite da tela e sobre ele, um marcador no formato de um pequeno quadrado amarelo onde deveria ser a posição inicial no espaço XY do objeto O_i .

A Figura 5.17 mostra o fluxograma (Pressman, 1995) do algoritmo de promediação fuzzy.

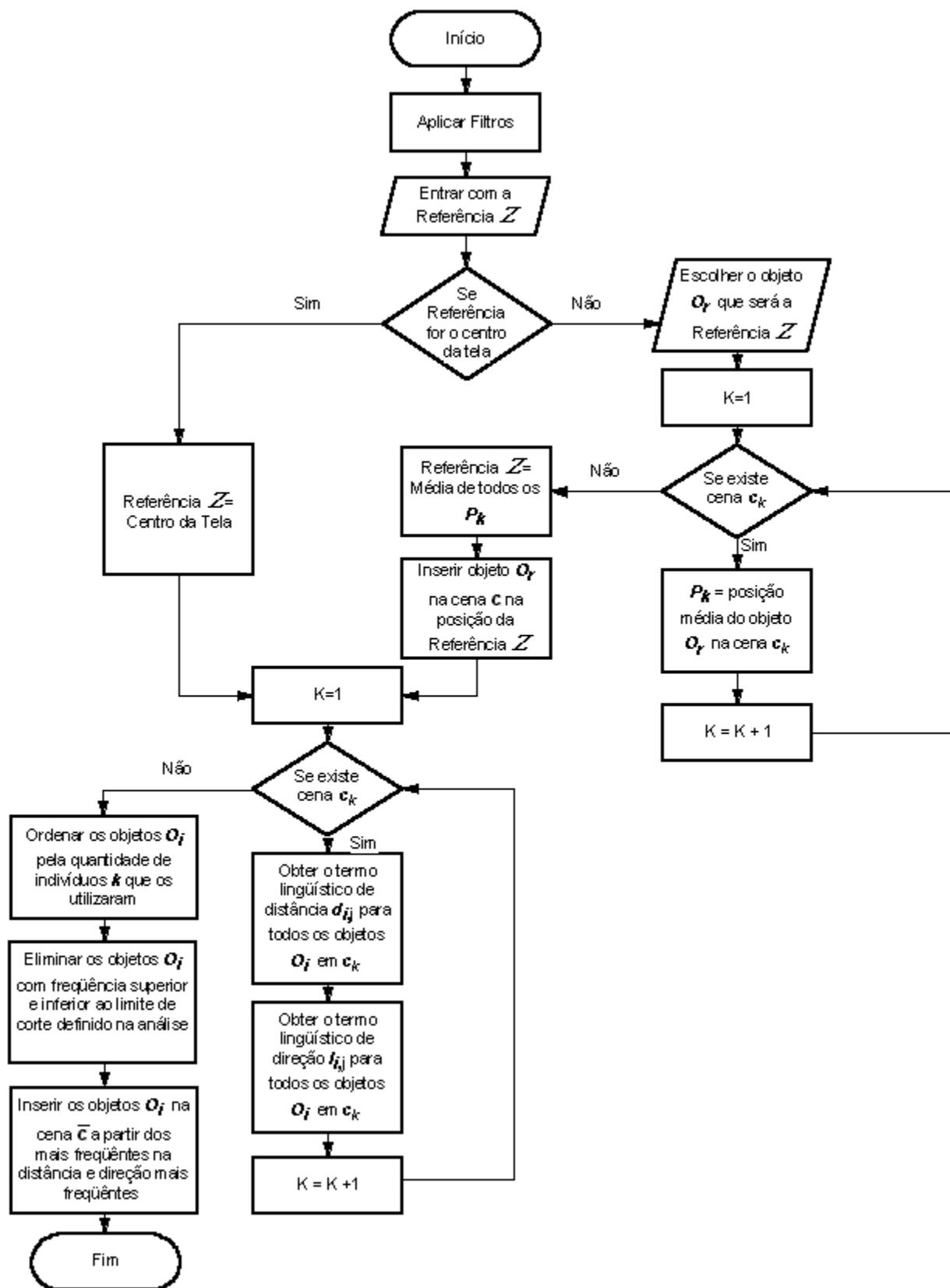


Figura 5.17 - Fluxograma do algoritmo de promediação fuzzy.

5.3 Critérios de aceitação de \bar{c}

A adequação de \bar{c} será validada pela sua comparação com uma cena \bar{c}_m obtida pela promediação de posições dos objetos utilizados em um conjunto \mathbf{C} formado por cenas semelhantes c_k , e principalmente por produzir ou não um conjunto visualmente aceitável por parte de qualquer pessoa.

A princípio, uma cena é visualmente aceitável se os objetos que a compõe seguem certas regras gerais de posicionamento, como por exemplo, em uma cena visualmente aceitável, uma casa não pode estar localizada dentro de um lago, ou no céu. Assim como uma casa não pode estar localizada em cima de outra.

Desta forma, pode ser observado que a figura de fundo de uma cena, possui um importante papel na definição das regras de posicionamento que um indivíduo usa para definir se uma cena é ou não visualmente aceitável. Numa cena que possua como figura de fundo uma paisagem, se houver um objeto casa, carro, ou pessoa, localizada na região referente ao céu desta paisagem, este objeto estará infringindo a regra de posicionamento, tornando esta cena visualmente inaceitável.

O fato do desenho de uma figura de fundo ser plano ou em perspectiva, também influencia de forma determinante nas regras de posicionamento. Em uma figura de fundo, cujo desenho esteja em perspectiva, um objeto que estiver localizado sobre outro, pode ser aceito pelas regras de posicionamento, já que podemos interpretar que um objeto está na frente ou atrás de outro. Dependendo do tipo do objeto, também é possível este tipo de interpretação em figuras de fundo que possuam desenho plano.

Outro fator que deve ser analisado na definição de regras de posicionamento é a ordem em que determinados objetos se encontram.

A Figura 5.18 – A, mostra um exemplo de uma cena em que o objeto “cerca” sobrepõe o objeto “casa” na região do telhado, dando a impressão que a cena está incoerente. Já a Figura 5.18 – B, mostra a mesma cena da Figura 5.18 – A, porém o objeto “casa” é que sobrepõe o objeto “cerca”, dando a clara impressão que a “casa” encontra-se à frente da “cerca”, respeitando assim as regras de posicionamento, pelo fato da figura de fundo (montanhas de neve e céu) ser um desenho em perspectiva assim como o objeto “casa”.

Às vezes, somente alterando a ordem (trazendo para frente ou enviando para trás) de determinados objetos, podemos transformar uma cena aparentemente incoerente numa cena que respeite totalmente as regras de posicionamento.

Desta forma, sempre que a sobreposição de objetos cause uma impressão de incoerência, mas que isso possa ser satisfatoriamente resolvido pela simples mudança na ordem de colocação destes objetos, podemos considerar que esta cena é considerada aceitável pelas regras de posicionamento.

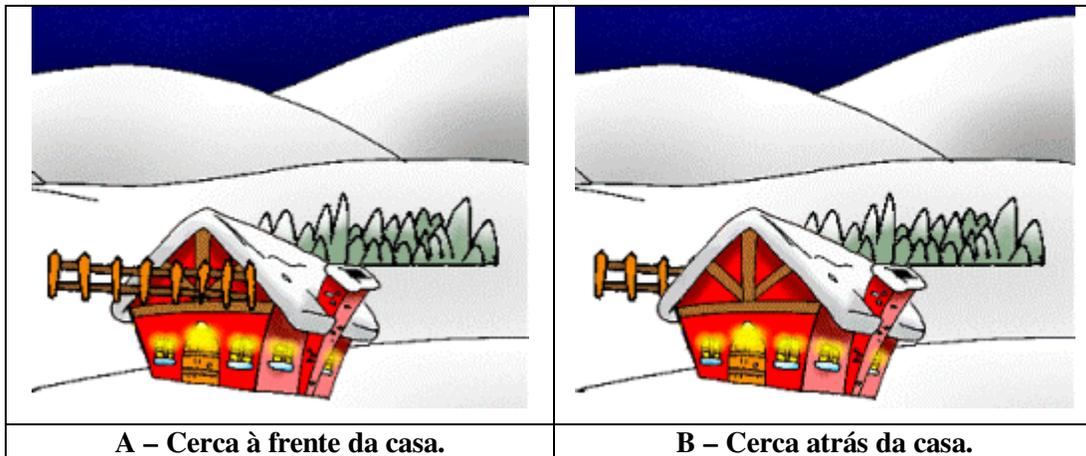


Figura 5.18 - Imagens com os mesmos objetos, porém com as ordens invertidas.

5.4 Resumo do Capítulo

Neste capítulo, foram mostrados dois algoritmos:

- Um para geração da cena média \bar{c}_m , aqui chamado de Promediação Convencional;
- E outro para geração da cena prototípica média \bar{c} , utilizando proposta da TCP, a qual chamamos Promediação Fuzzy.

No capítulo a seguir, será mostrado como é feita a análise das descrições das cenas desenvolvidas pelos voluntários.

6 Análise das Descrições l_k das Cenas c_k

A análise das descrições l_k das cenas c_k criadas pelos voluntários foi feita a partir da interação do **VMC** com o sistema Jargão, desenvolvido por Rocha e colaboradores (Rocha et al., 1992, Guilherme, 1996).

O **VMC** é responsável por juntar todas as descrições l_k e pedir os parâmetros necessários para o processamento. Feito isso, estas informações são passadas para o sistema Jargão, que é um software externo ao **VMC** e a partir daí as descrições l_k são processadas e os resultados devolvidos ao **VMC**.

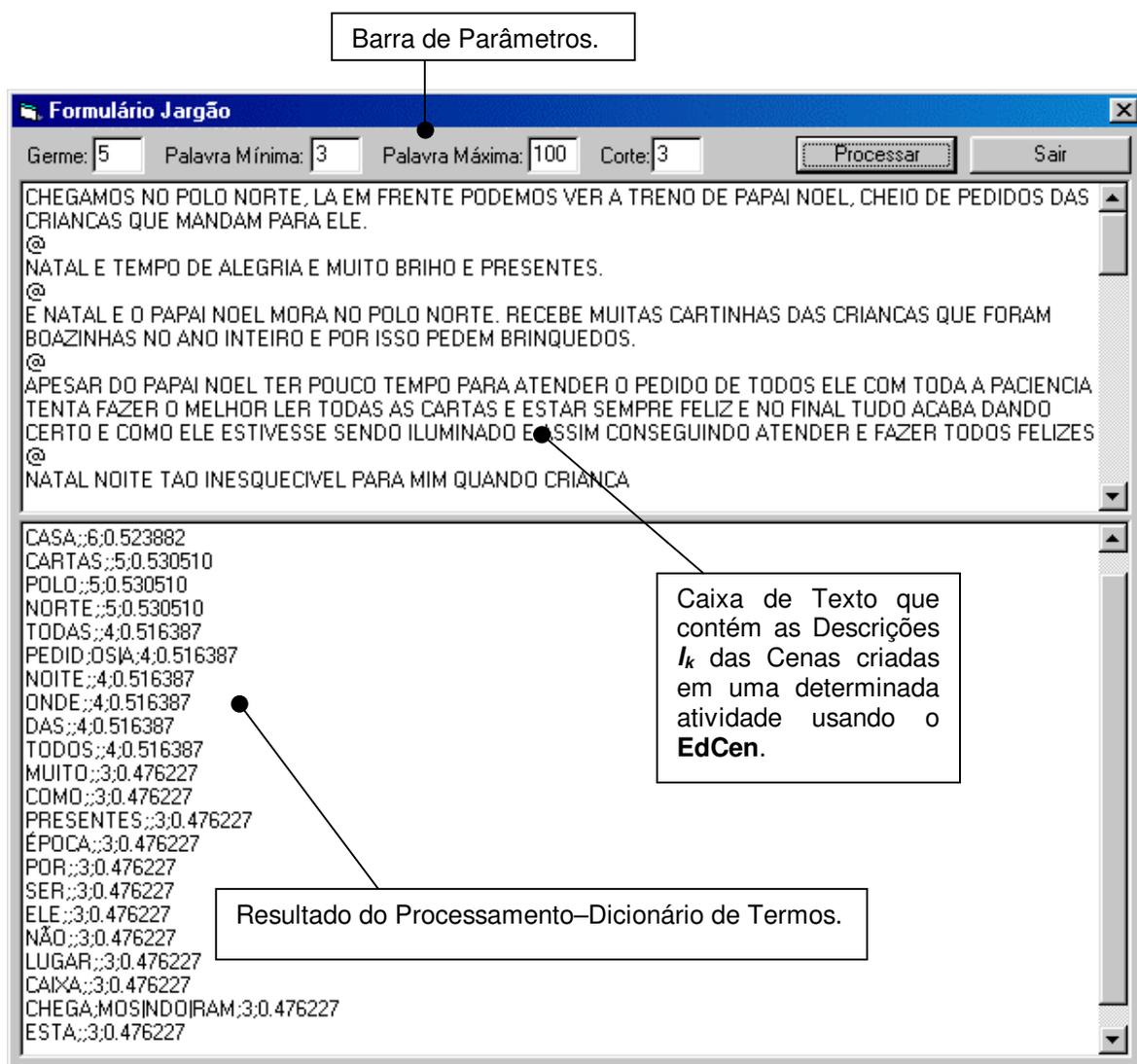


Figura 6.1 - Analisador de Textos.

Ao ser ativado, através do botão “Textos” no **VMC** (Figura 4.6), o **Analisador de Textos** carrega todas as descrições I_k das cenas c_k semelhantes (Figura 6.1), onde as descrições são separadas umas das outras pelo caractere “@”.

Estando o **Analisador de Textos** ativo, deve-se então atribuir os valores desejados aos parâmetros (barra de parâmetros - Figura 6.1), que serão passados ao sistema Jargão para o processamento. Este primeiro processamento terá como resultado o Dicionário de Termos (Figura 6.1).

Portanto, o Dicionário de Termos (Figura 6.1), é obtido do conjunto de descrições I_k , a partir da filtragem de todas palavras que obedeçam aos parâmetros estruturais definidos. Esses parâmetros são os seguintes:

- <Germe>: É um número inteiro que indicará o tamanho do radical que melhor descreve cada palavra.
Ex: 5 -> (Banana // Germe: Banan)
- <Palavra Mínima>: Número inteiro, que informa ao sistema para só realizar a extração das palavras da base de dados que tiverem o número de caracteres maior ou igual ao "Min".
Ex: 3 -> (So // esta palavra não entra na análise)
3 -> (Sol // esta palavra entra na análise)
- <Palavra Máxima>: Número inteiro, que informa ao sistema para só realizar a extração das palavras da base de dados que forem menores ou iguais ao "Max".
Ex: 9 -> (testemunha // esta palavra é desconsiderada)
9 -> (testar // esta palavra entra na análise)
- <Corte>: Número inteiro. Corta as palavras que tiverem uma freqüência menor ou igual a "corte".
Ex: 2 -> (Freqüência: 2, desconsidera a palavra)
2 -> (Freqüência: 3, considera palavra para análise)

A Figura 6.2, mostra uma parte dos Dicionários de Termos das atividades Cena mais Engraçada e Fábrica de Brinquedos. Podemos observar que o Dicionário de Termos, mostra entre outras coisas, as palavras encontradas, e a quantidade de ocorrência destas. O sistema Jargão de acordo com o valor passado no parâmetro germe, é capaz de analisar diferentes terminações de um mesmo radical e classificá-los como uma mesma palavra, como por exemplo: **ENTRA; DA | NDO | R | ; 6**, ou seja, foram encontradas seis ocorrências do radical “ENTRA”, com as seguintes terminações: “DA”, “NDO”, “R”.

MULHER; ; 11; BIGODE; ; 10; ENTRA; DA NDO R ; 6; MÃE; ; 5; MANADA; ; 5; BRIGA; NDO; 5; CONFU; SAO SÃO NDE ; 5; CAMIS; A ETA; 4; ROSTO; ; 3; VASCO; ; 3; VASCAINO; ; 3;	NORTE; ; 5; PEDID; OS A; 4; NOITE; ; 4; PRESENTES; ; 3; ÉPOCA; ; 3; SER; ; 3; ELE; ; 3; NÃO; ; 3; LUGAR; ; 3; CAIXA; ; 3; CHEGA; MOS NDO RAM; 3;
DICIONÁRIO DE TERMOS – Cena mais Engraçada	DICIONÁRIO DE TERMOS – Fábrica de Brinquedos

Figura 6.2 - Exemplo de Dicionário de Termos gerados para as atividades Cena mais Engraçada e Fábrica de Brinquedos

A partir das palavras mais frequentes encontradas no Dicionário de Termos, é montado o Dicionário de Sintaxe, expresso na forma de uma gramática **G** e contém palavras que espera-se encontrar nas descrições I_k . De acordo com o resultado da análise do Dicionário de Sintaxe, feito pelo sistema Jargão, pode-se validar ou não este dicionário.

A gramática **G** (Figura 6.3) usada pelo Jargão para análise de frases, descreve os tipos de frase a serem aceitas como conjuntos ordenados de palavras que definem um conceito de interesse a ser investigado na base de textos.

BIGODE/[MULHER DE BIGODE] (-5, 5, 1, 1) AAA& //MULHER_MOÇA (-5, 5, 1) aaa& \$ENTR_CHEG/[ENTRADA DO VASCAINO] (-5, 5, 1, 1) AAA& //VASC (-5, 5, 1) aaa& \$CHAM_ACLAM_ENCONT_RECON_CONFUND/[MÃE] (-5, 5, 1, 1) AAA& //MÃE (-5, 5, 1) aaa& \$ENTR_CORR/[MANADA] (-5, 5, 1, 1) AAA& //MANADA (-5, 5, 1) aaa&	CASA_MORA/[MORADIA DO PAPAÍ NOEL] (-5, 5, 2, 2) AAA& //NOEL (-5, 5, 1) aaa&BBB& //NORTE (-5, 5, 1) bbb& \$PEDI_CARTA_PRESEN/[PRESENTES] (-5, 5, 1, 2) AAA& //NOEL (-5, 5, 1) aaa&BBB& //Grande_Muit_Bastante (-5, 5, 1) bbb& \$ACREDIT_BRILHO_ILUMI_PERDOA_FELI/[EMOCIONAL] (-10, 10, 1, 1) AAA& //NATAL_PAPAÍ NOEL (-5, 5, 1) aaa&
DICIONÁRIO DE SINTAXE-Cena mais Engraçada	DICIONÁRIO DE SINTAXE-Fábrica de Brinquedos

Figura 6.3 - O dicionário de sintaxe.

As palavras de interesse são definidas pelo seu radical, por exemplo: ENTR é um radical associado às palavras ENTRAR, ENTRADA, ENTROU, etc. A sinonímia de palavras é declarada da seguinte forma:

Palavra1_Palavra2_Palavra3_....._PalavraN

Exemplo: CASA_MORA

A vantagem de se usar sinônimos é, portanto, o fato do sistema Jargão aceitar diferentes palavras como símbolos iniciais, não-terminais, ou terminais. No exemplo mostrado anteriormente: **CASA_MORA**, o sistema Jargão aceitará frases que contenham tanto o radical **CASA** quanto o radical **MORA**.

Toda a busca no texto para identificação das possíveis frases aceitas pela gramática **G** é feita a partir da identificação inicial de palavras chaves, e depois de seus complementos ou palavras secundárias dentro de uma vizinhança permitida. Essa busca é realizada da seguinte maneira:

- Identifica-se o código no radical de uma palavra chave (por exemplo: "CASA_MORA"), que são 3 letras maiúsculas terminada em "&" (por exemplo: "AAA&") e identifica-se um conceito definido entre colchetes (por exemplo: "[MORADIA DO PAPAÍ NOEL]");
- Encontrado o código chave, o sistema vai à procura de uma associação de palavras definidas por códigos do tipo "aaa&XXX&", onde os caracteres "aaa" significam que esta é uma palavra secundária (por exemplo: NOEL aaa&BBB&) e pode ser associada ao código chave "AAA". O segundo código "XXX" desses radicais secundários, define agora a nova classe de palavras a ser buscada. Desta forma, montam-se níveis hierárquicos de procura de palavras num texto.
- O grau de confiança na ocorrência do conceito por um conjunto de regras (por exemplo: AAA&; aaa&BBB&; bbb&) é calculado através da relação do número de regras obedecidas pelo número total de regras da cadeia de derivação (por exemplo: AAA& → aaa&BBB& → bbb&);

Assim como proposto pela equação (XVIII), os códigos do tipo AAA& definem o conjunto de símbolos iniciais **S**. Já o código aaa&BBB& define símbolos intermediários **N** e ccc& define um símbolo terminal **T**. As regras de escrita dependem apenas do relacionamento entre caracteres desses códigos, pois o pareamento é permitido entre as maiúsculas e as minúsculas do mesmo caractere.

Uma frase \bar{I} , é portanto, uma frase que deve ser aceita pela gramática definida pelo dicionário de sintaxe, tendo um grau de pertinência acima de 0.5 (mais de 50%), de modo que, também será composta pelos termos mais freqüentes, já que estes compõe o dicionário de sintaxe.

Além disso, a frase \bar{I} gerada, deve produzir uma descrição aceitável por parte de qualquer pessoa. A análise para a aceitação ou não de \bar{I} é feita de forma subjetiva, onde uma descrição pode ser considerada aceitável desde que ela não desrespeite certas regras de correlação verbal / visual.

As regras de correlação verbal / visual, são as regras que avaliam se \bar{I} pode ser uma descrição aceitável para uma cena \bar{c} . Essas regras, que são baseadas no bom senso, não permitem que uma cena \bar{c} , tenha uma descrição \bar{I} sobre objetos que não façam parte de \bar{c} . Por exemplo, uma cena \bar{c} sobre o tema natalino, composta pelos elementos pictóricos mostrados na Figura 4.3, não permite uma descrição \bar{I} que fale entre outras coisas, sobre objetos referentes ao carnaval, páscoa, ou qualquer outro assunto não relacionado ao tema da cena \bar{c} .

Além dos comentários sobre os objetos de uma cena \bar{c} , uma descrição \bar{I} também pode conter comentários sobre o aspecto emocional que esta cena proporciona. Foi comum, por exemplo, nas descrições I_k das cenas c_k da atividade “Fábrica de Brinquedos”, a ocorrência de citações sobre perdão, felicidade, o brilho do natal, etc.

As descrições emocionais, também devem ser aprovadas pelas regras de correlação verbal / visual, para que uma descrição possa ser considerada aceitável. Mesmo as emoções sendo algo particular de cada indivíduo, não podemos por exemplo, aceitar em uma descrição \bar{I} sobre a cena mais engraçada, comentários sobre fatos da cena que provocam tristezas, frustrações, etc, ou seja, uma descrição emocional deve ser coerente ao tema abordado na cena \bar{c} .

6.1 Resumo do Capítulo

Neste capítulo, foi mostrado o processo de análise, as descrições I_k das cenas c_k criadas pelos voluntários k e a interação do **VMC** com o sistema Jargão.

No próximo capítulo, serão apresentados os resultados obtidos com a pesquisa.

7 Resultados

Como descrito anteriormente no tópico 4.1 - Grupo Experimental, um grupo de 23 adultos, com nível de escolaridade médio ou superior desenvolveu as seguintes tarefas:

- Ouviram um texto gravado de uma história de humor escrita por Luis Fernando Veríssimo, mostrado na Tabela 4.1;
- Ouviram novamente o trecho inicial da história, marcado em azul na Tabela 4.1 e compuseram uma cena sobre o trecho ouvido;
- Compuseram uma cena sobre o episódio da história considerado mais engraçado;
- Desenvolveram uma cena sobre o tema natalino Fábrica de Brinquedos.
- Ao término da composição das cenas do episódio mais engraçado e da fábrica de brinquedos, os voluntários fizeram uma descrição escrita das cenas montadas.

Após a aquisição dos dados, foi utilizado o **VMC** para o cálculo da cena média \bar{c}_m e da cena prototípica média \bar{c} , utilizando os algoritmos apresentados no capítulo 5 (Geração de cenas prototípicas médias), o de Promediação e o de Promediação Fuzzy, que utiliza a técnica da **TCP** (Zadeh,2001).

As cenas médias \bar{c}_m que apresentaram como resultado uma cena incoerente, com objetos fluando ou sobrepostos, portanto, que desrespeitaram as regras de posicionamento, foram corrigidas por um ser humano, utilizando o **EdCen**, com o intuito de obter uma cena visualmente coerente sobre o ponto de vista humano. Estas correções são apenas algumas alterações no posicionamento dos objetos a partir da cena média \bar{c}_m original, a fim de que todos os objetos respeitassem as regras de posicionamento. Estas cenas corrigidas, que podemos considerar como cenas visualmente aceitáveis, serão comparadas com suas respectivas cenas médias \bar{c}_m .

Para as tarefas em que foram pedidas descrições I_k das cenas c_k montadas, usou-se o **VMC** para análise destas descrições I_k .

As funções de pertinência usadas para a codificação da distância euclidiana $d_{i,j}$ (entre um objeto O_i e a referência Z) para uma variável lingüística $d_{i,j}$, e as funções de pertinência usadas para a recodificação da distância referenciada $d_{i,j}$ para uma distância euclidiana $d_{i,j}$ nas atividades “Abertura”, “Cena Mais Engraçada” e “Fábrica de Brinquedos” são iguais às mostradas nos tópicos 5.2.2 e 5.2.4.

Os resultados obtidos serão mostrados segundo a atividade executada.

7.1 Abertura

Na atividade “Abertura” foi disponibilizado aos voluntários uma série de elementos pictóricos (Figura 4.1), dos quais vários não faziam referência ao trecho ouvido (Tabela 4.1 em azul), trecho sobre o qual os voluntários deveriam montar uma cena c_k .

Como se pode observar na Figura 7.2 que mostra todos os elementos pictóricos usados e a quantidade k de indivíduos que os utilizaram, mais de 74% dos voluntários escolheram os elementos pictóricos que faziam referência ao recital e à entrada do vascaíno, sendo que apenas 26% dos voluntários utilizaram elementos pictóricos que faziam alusão a outros trechos do conto, sendo que dois elementos pictóricos O_i não chegaram a ser usados. Em razão disto, aplicou-se um limiar de frequência de uso, no qual só entraram no cálculo na cena média \bar{c}_m e da cena prototípica média \bar{c} , objetos O_i que foram usados por mais de 74% dos indivíduos.

7.1.1 Cenas médias \bar{c}_m geradas através da Promediação Convencional

A Figura 7.1, mostra a cena média \bar{c}_m original, gerada através do algoritmo de promediação e a cena média \bar{c}_m corrigida por um ser humano através do **EdGen**, que é gerada a partir da cena original. Algumas alterações no posicionamento dos objetos são necessários, a fim de deixar uma cena mais clara e coerente.

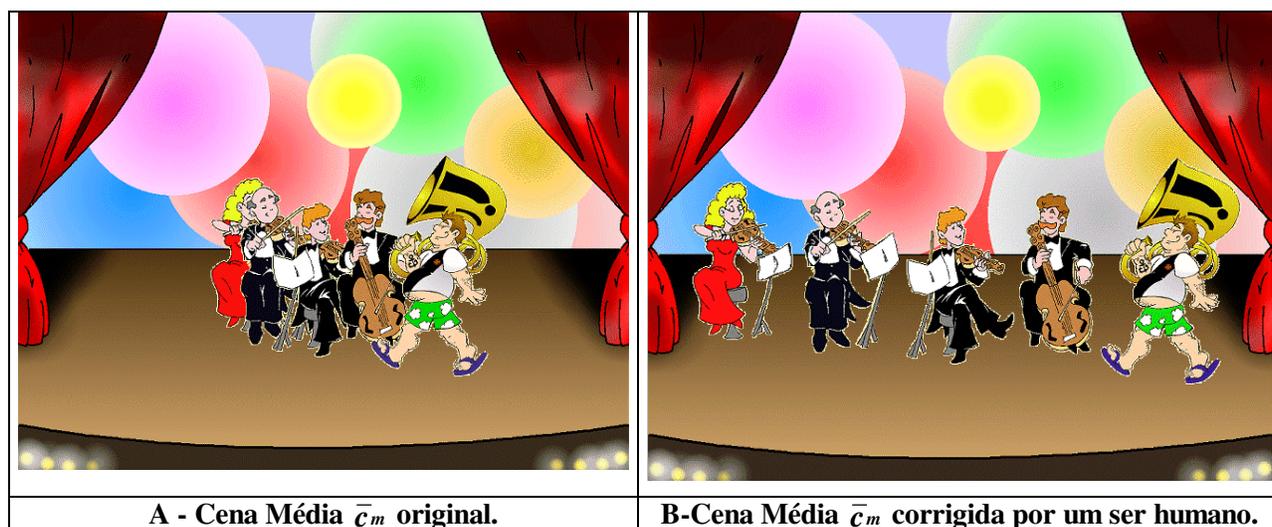


Figura 7.1 - Cena Média \bar{c}_m original e corrigida por um ser humano com frequência de uso de objetos O_i superior a 16 indivíduos.

A Figura 7.1 – A mostra a cena média \bar{c}_m gerada. Como se pode observar, a cena está bem montada, porém um pouco amontoada junto ao centro. Por isso na cena média \bar{c}_m corrigida, mostrada na Figura 7.1 – B, foram feitas as seguintes correções:

- Deslocamento do objeto “mulher violoncelista” para a esquerda;
- Leve deslocamento do objeto “violonista careca” para a esquerda;
- Leve deslocamento do objeto “homem violoncelista” para a direita;
- Deslocamento do objeto “vascaíno com a tuba” para a direita.

7.1.2 Cenas prototípicas médias \bar{c} geradas através de Promediação Fuzzy



Figura 7.2 - Tela de seleção do elemento pictórico O_r para ser tomado como referência Z na tarefa “Abertura”.

A Figura 7.3, mostra duas cenas prototípicas médias \bar{c} geradas, utilizando como referência Z o centro da tela e o elemento pictórico O_r “mulher violoncelista”, respectivamente.

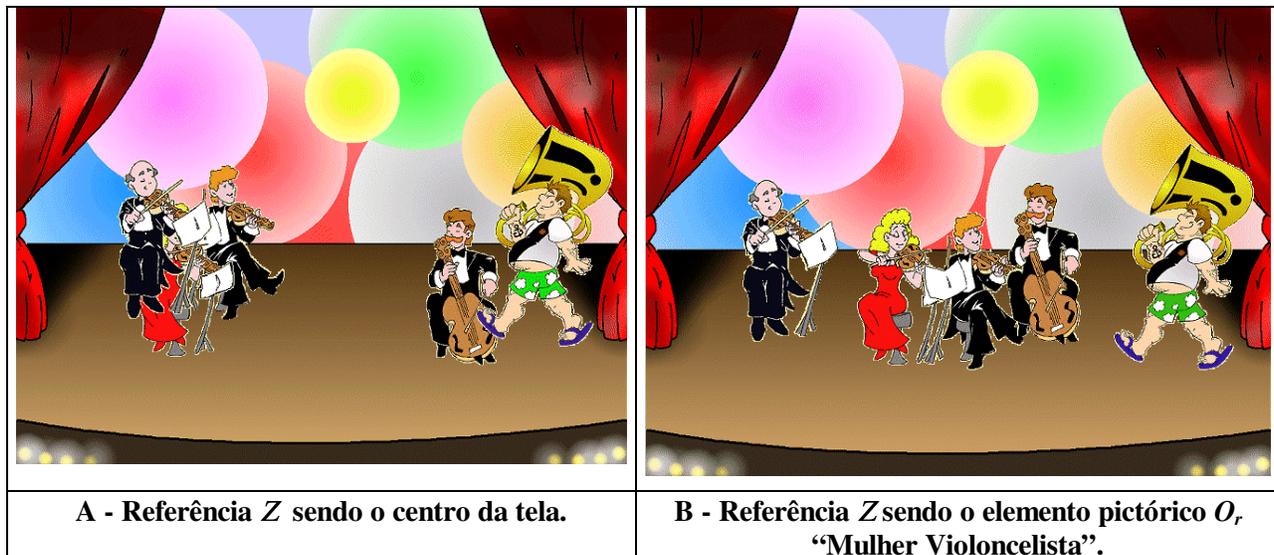


Figura 7.3 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Abertura”, sendo o limiar de corte acima de 16 objetos.

Como se pode observar, a referência Z no centro da tela não proporcionou uma boa imagem, já que houve sobreposição de objetos, infringindo as regras de posicionamento e gerando uma imagem incoerente. A melhor referência Z encontrada foi o elemento pictórico O_r “mulher violoncelista”. Com esta referência a cena gerada ficou clara e coerente, sem a necessidade de nenhuma correção estrutural.

Também foi possível utilizar como referência Z , os elementos pictóricos “violinista careca” e “violinista”, porém, as imagens geradas tiveram uma qualidade apenas regular, sendo que, alguns elementos pictóricos ficaram muito próximos uns dos outros deixando a cena um pouco rebuscada. Já com os elementos pictóricos “vascaíno com a tuba” e “homem violoncelista”, não foi possível obter imagens coerentes.

7.2 Cena Mais Engraçada

Devido à grande variabilidade de cenas consideradas engraçadas pelos voluntários, não foi possível gerar uma cena média \bar{c}_m ou cena prototípica média \bar{c} que utilizasse todos os elementos pictóricos, já que em todas as tentativas a cena ficava extremamente carregada, não sendo possível uma interpretação do cenário montado ou então uma coerência de idéias.

Aplicou-se um limiar de corte de frequência de uso dos objetos O_i por voluntários k , dividindo-se o grupo inicial em dois subgrupos:

- um com frequência de uso igual ou superior a onze elementos pictóricos, ou seja, só entrou no cálculo da cena média \bar{c}_m e da cena prototípica média \bar{c} os objetos O_i utilizados por mais de 11 voluntários k , o que representa aproximadamente 48% dos voluntários que executaram esta tarefa.
- e outro subgrupo com frequência de uso entre cinco e dez objetos.

O primeiro subgrupo, que possui os objetos O_i mais usados, faz referência a dois trechos específicos do conto “O Recital”. Um quando o vascaíno olha para a violoncelista que está com o bigode e grita “Mamãe!”, e o outro quando entra no palco a manada de zebus. Isso significa que a maioria dos indivíduos que participaram da pesquisa considerou estas passagens do conto como as mais engraçadas.

Já o segundo subgrupo, faz alusão à entrada do vascaíno com a tuba e às discussões.

7.2.1 Cenas médias \bar{c}_m geradas através da Promediação Convencional

As cenas médias \bar{c}_m mostradas na Figura 7.4 são a original e a corrigida da tarefa “Cena Mais Engraçada” geradas a partir do primeiro subgrupo, ou seja, apenas com os objetos O_i usados por mais de 11 voluntários.

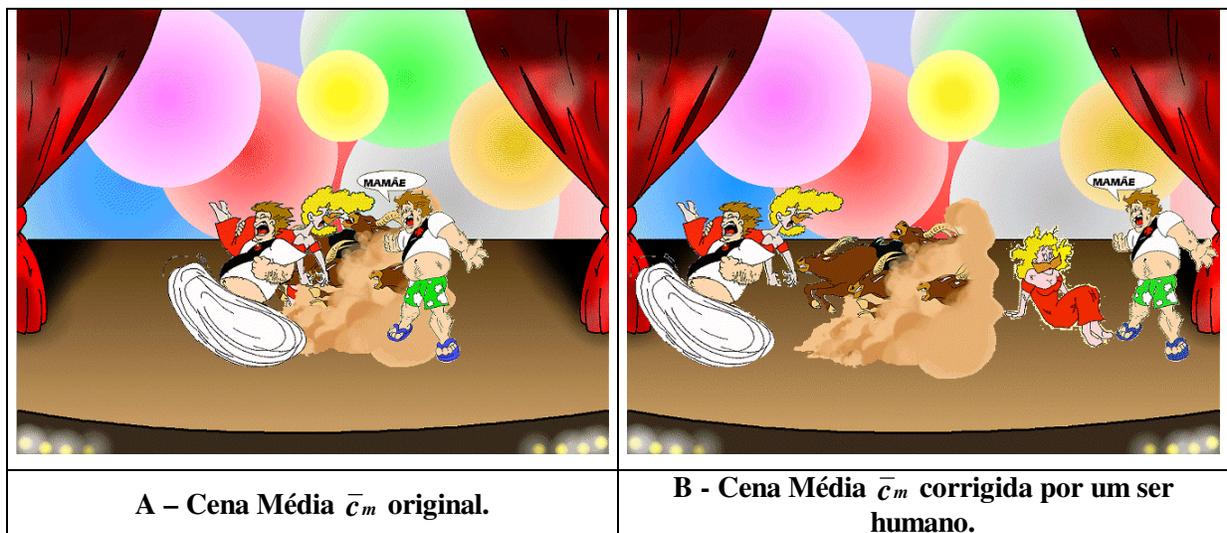


Figura 7.4 - Cena Média \bar{c}_m original e corrigida por um ser humano com frequência de uso de objetos O_i superior a 11 indivíduos.

Já a cena corrigida foi obtida a partir da cena média \bar{c}_m , mas com as seguintes alterações:

- Deslocamento dos objetos “vascaíno carregando a violoncelista” e “a manada de zebus” para a esquerda;
- Deslocamento do objeto “vascaíno gritando” para a direita;
- Leve deslocamento do objeto “mulher caída” para a direita;
- Deslocamento do objeto “bigode” sobre a posição referente ao rosto do objeto “mulher caída”.

Verifica-se, portanto, que foi necessária a reordenação do posicionamento dos objetos O_i da cena original para que estes não ferissem as regras de posicionamento, havendo desta forma coerência no posicionamento dos objetos, possibilitando um entendimento da cena média \bar{c}_m gerada. Sem estas correções, a cena média \bar{c}_m gerada fica incoerente.

A Figura 7.5, mostra a cena média \bar{c}_m original e corrigida do segundo subgrupo, ou seja, o grupo com frequência de uso de objetos O_i , por no mínimo cinco e máximo de 10 indivíduos.

A cena média \bar{c}_m gerada no segundo subgrupo é um pouco mais ordenada que a do primeiro subgrupo, mas mesmo assim necessitou de uma correção para que se adequasse às regras de posicionamento. Essa correção foi o deslocamento do tubista para a direita.

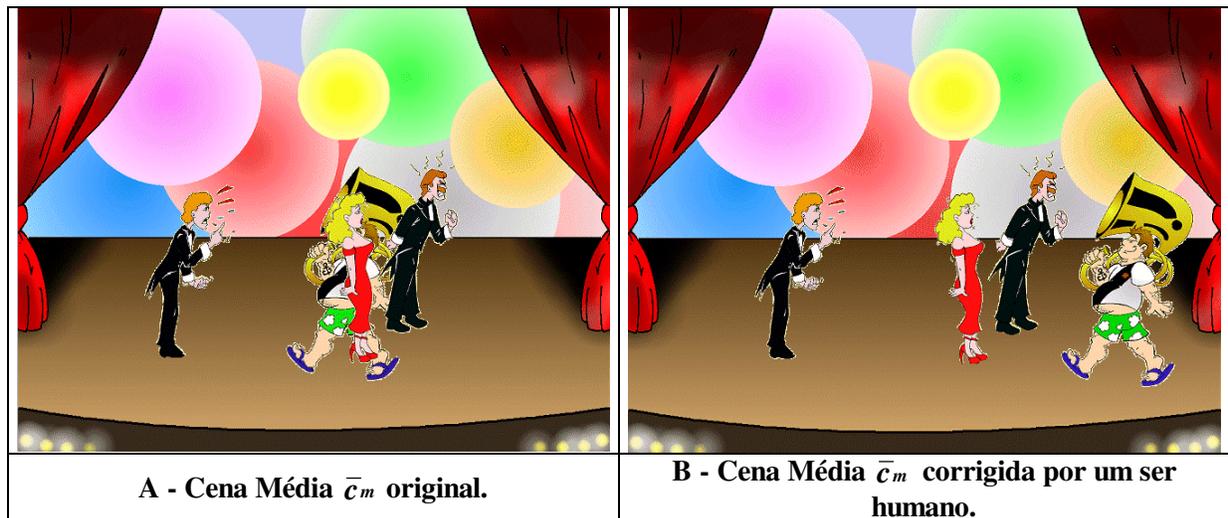


Figura 7.5 - Cena Média \bar{c}_m original e corrigida por um ser humano com frequência de uso de objetos O_i , por no mínimo cinco e máximo de 10 indivíduos.

7.2.2 Cenas prototípicas médias \bar{c} geradas através de Promediação Fuzzy



Figura 7.6 -Tela de seleção do elemento pictórico O_i para ser tomado como referência Z na tarefa “Cena Mais Engraçada”.

A Figura 7.6, mostra os elementos pictóricos O_i disponíveis para serem usados como referência Z e suas respectivas quantidades de uso. A Figura 7.6 desconsidera o uso de filtros de frequência, mostrando todos os elementos pictóricos que foram usados no conjunto C de cenas semelhantes c_k .

As cenas prototípicas médias \bar{c} mostradas na Figura 7.7 são as cenas geradas a partir do primeiro subgrupo, o qual possui apenas os objetos O_i usados por mais de 11 voluntários.

Estudou-se a composição da cena prototípica média \bar{c} , fixando a referência Z em vários objetos O_i e também no centro da tela (Figura 7.7).

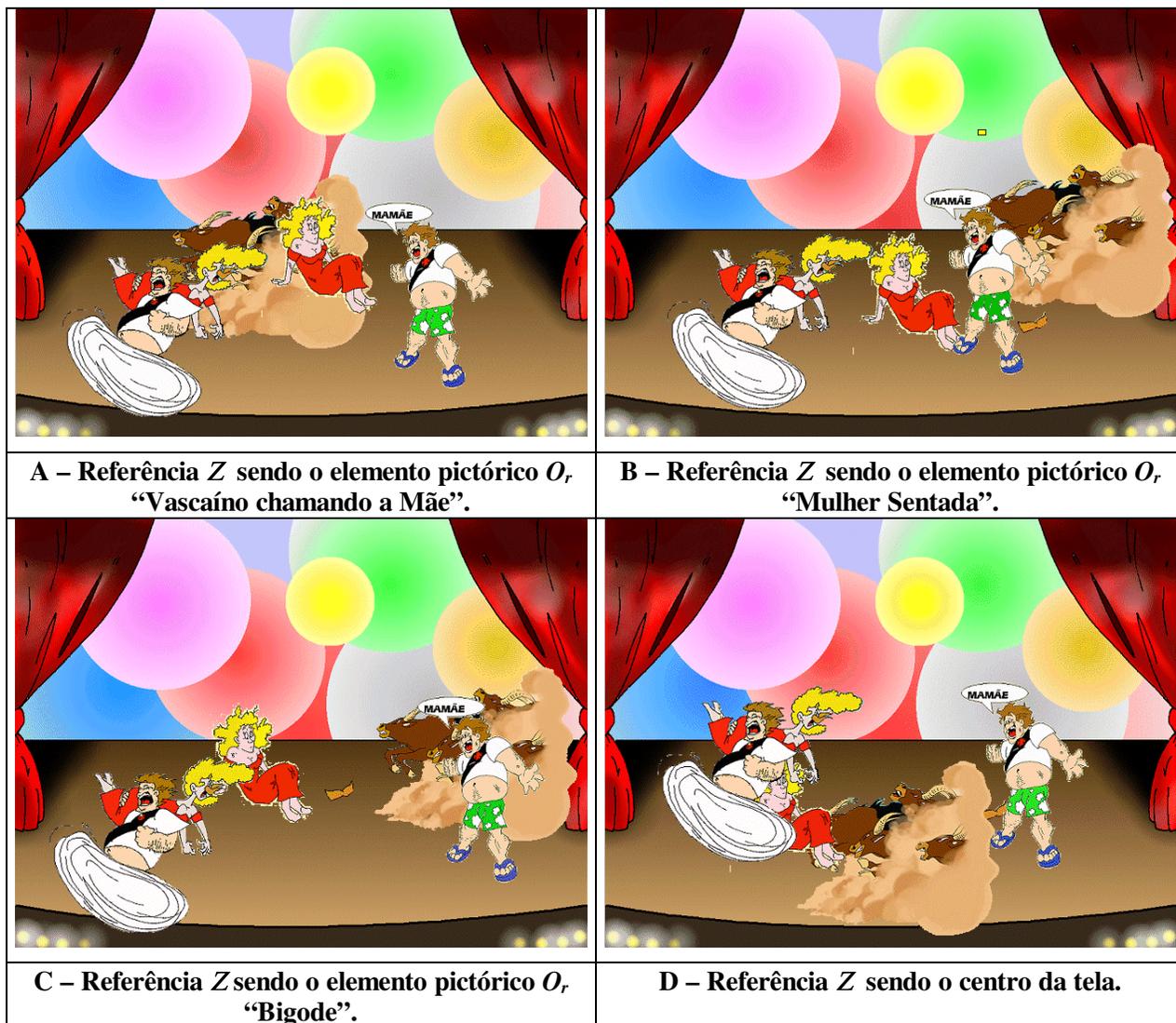


Figura 7.7 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Cena Mais Engraçada”, sendo o limiar de corte acima de 11 objetos.

As cenas prototípicas médias \bar{c} montadas, tendo como referência Z os elementos pictóricos “vascaíno chamando a violoncelista de mamãe”, Figura 7.7 – A, “mulher sentada”, Figura 7.7 – B, e o “bigode”, Figura 7.7 – C, são exemplos de cenas geradas que possuem um alto grau de clareza e facilidade de entendimento, melhor do que a da cena média \bar{c}_m gerada a partir de uma promediação convencional, visto que a primeira, não necessita de retoques.

Já a cena prototípica média \bar{c} , gerada com a referência Z sendo o centro da tela Figura 7.7 – D, apresentou objetos sobrepostos, o que tornou a cena incoerente.

Na maior parte das cenas montadas, o elemento pictórico “bigode” foi colocado sobre o rosto do elemento pictórico “mulher sentada”, porém, no cálculo da cena prototípica média \bar{c} , o bigode nunca ficou sobre o rosto da mulher.

Já as cenas prototípicas médias \bar{c} mostradas na Figura 7.8 são as cenas geradas a partir do segundo subgrupo, o qual possui apenas os objetos O_i usados por mais de cinco e menos de 10 voluntários. As Figura 7.8 A e B mostram respectivamente duas cenas prototípicas médias \bar{c} geradas, sendo que a referência Z da primeira é o centro da tela e a da segunda, é o elemento pictórico O_i “mulher violoncelista”.

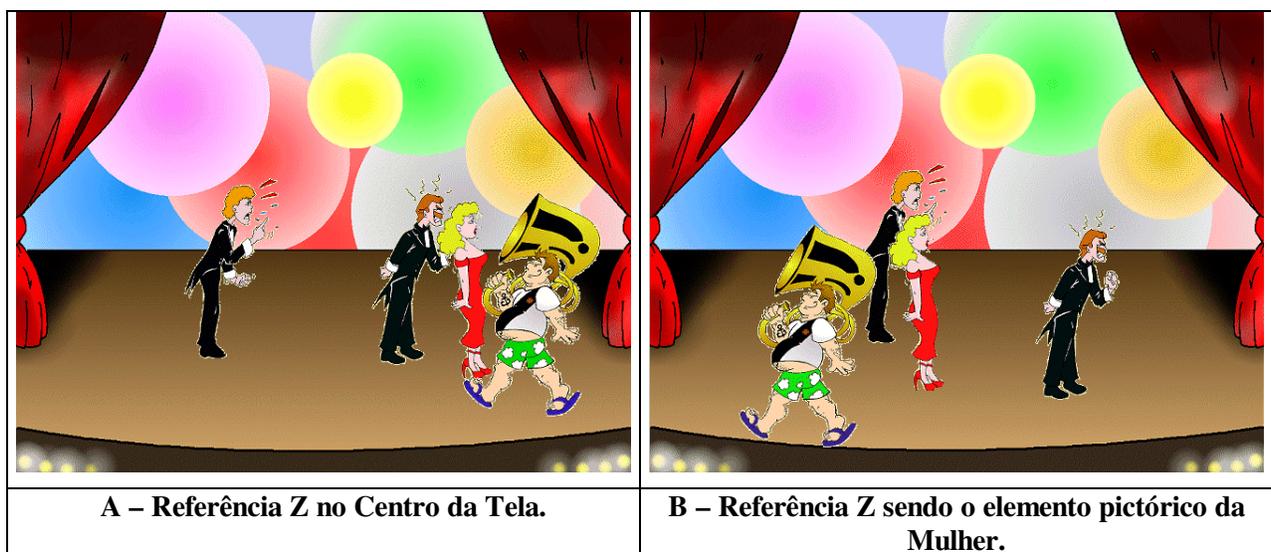


Figura 7.8 - Cenas prototípicas médias com limiar de corte acima de 5 objetos e abaixo de 10 objetos.

7.2.3 Análise das Descrições l_k das Cenas c_k

Apesar das cenas terem sido compostas por 23 voluntários, quatro deles recusaram-se a fazer a descrição da cena montada, somando-se assim dezenove descrições l_k criadas.

O dicionário de sintaxe, apresentado na Tabela 7.1, foi montado a partir da análise dos termos mais usados (Capítulo 6) pelos voluntários para a descrição desta atividade.

Vale observar que neste dicionário de sintaxe foram utilizados sinônimos nas buscas feitas aos textos que contêm as descrições, como por exemplo: **ENTR_CHEG**, é um símbolo inicial **S** que busca como complemento um símbolo terminal **T** com a seqüência de caracteres **VASC**. Isso foi feito porque várias pessoas fizeram menção à entrada do vascaíno, porém, utilizando frases do tipo: “entrada do

vascaíno no palco”, “chegada do vascaíno”, “entrou um vascaíno”, dentre outras. Desta forma, utilizamos como símbolo inicial apenas os radicais **ENTR_CHEG** para permitir diferentes conjugações dos verbos e usamos o terminal **VASC** para permitir que fosse encontradas também palavras como “Vasco” e “Vascaíno”, escrito com e sem acento agudo no caractere “i”. A vantagem de se usar sinônimos é, portanto, o fato do sistema Jargão aceitar diferentes palavras como símbolos iniciais, não-terminais, ou terminais. No caso mostrado anteriormente, onde o símbolo inicial é **ENTR_CHEG** e o terminal é **VASC**, o sistema Jargão aceitará frases que contenham tanto o radical **ENTR** quanto **CHEG**, desde que possuam o símbolo terminal **VASC** dentro do intervalo definido. Este intervalo é definido pelos números **-5** e **5**, que estão dentro do parênteses e indica que achado um símbolo inicial deve ser procurado seu complemento entre as cinco palavras anteriores ou posteriores ao símbolo inicial.

CONCEITOS	
ENTR_CHEG/[ENTRADA DO VASCAINO](-5,5,1,1)AAA& //VASC (-5,5,1)aaa& \$ Freqüência de ocorrência: 9 em 19 textos	
BIGODE/[MULHER DE BIGODE](-5,5,1,1)AAA& //MULHER_MOÇA (-5,5,1)aaa& \$ Freqüência de ocorrência: 6 em 19 textos	
CHAM_ACLAM_ENCONT_RECON_CONFUND/[MÃE](-5,5,1,1)AAA& //MÃE (-5,5,1)aaa& \$ Freqüência de ocorrência: 7 em 19 textos	
ENTR_CORR/[MANADA](-5,5,1,1)AAA& //MANADA (-5,5,1)aaa& \$ Freqüência de ocorrência: 10 em 19 textos	
ELEMENTO PICTÓRICO	FREQÜÊNCIA
Vascaíno Rasgando Partitura	0
Manada	13
Bigode	11
Mulher Depois da Briga	13
Violinista Brigando	5
Vascaíno Brigando	2
Músico Brigando	5
Vascaíno Chamando Mamãe	14
Vascaíno Levando Mulher	15
Vascaíno com Tuba	8
Vascaíno Sedutor	1
Mulher Assustada	8

Tabela 7.1 - Descrição da Cena Mais Engraçada.

Outro exemplo de sinonímia usada foi **CHAM_ACLAM_ENCONT_RECON_CONFUND**, capaz de representar as conjugações dos verbos “chamar”, “aclamar”, “encontrar”, “reconhecer” e “confundir”.

O resultado da análise das descrições I_k , associado à cena prototípica média \bar{c} como vimos na Figura 7.7 e Figura 7.8, é mostrado na Tabela 7.1. O relato sobre a entrada do vascaíno no palco, aparece em 9 dos 19 textos analisados e o elemento pictórico correspondente (Vascaíno com a Tuba) foi incluído em 8 dos 19 desenhos processados. A referência à mulher de bigode foi feita por 6 dos voluntários na criação dos textos, enquanto que a composição do objeto O_i “mulher caída com bigode” apareceu em 11 dos 19 desenhos. Em 14 das cenas foi incluído o objeto O_i “vascaíno chamando a violoncelista de mãe”, porém a referência no texto a esse conceito aparece em apenas 7 dessas cenas. Finalmente, os objetos O_i “vascaíno carregando a mulher” e “entrada da manada de zebus” foram utilizados por 13 voluntários, enquanto que referências a esses fatos apareceram em 10 dos 19 textos.

7.3 Fábrica de Brinquedos

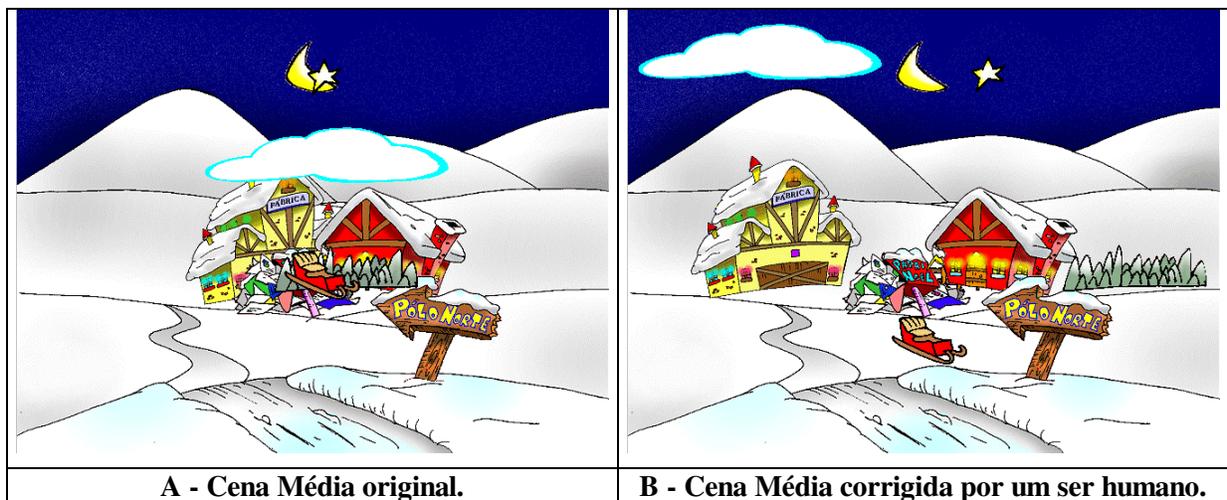
Na atividade “Fábrica de Brinquedos”, foram geradas cenas médias \bar{c}_m e prototípicas médias \bar{c} , em que foram usados todos os elementos pictóricos e a filtragem de frequência de uso.

A frequência mínima de uso foi fixada em onze elementos pictóricos, ou seja, apenas entraram no cálculo da cena média \bar{c}_m e da cena prototípica média \bar{c} os objetos O_i utilizados por mais de 11 voluntários k , o que representa aproximadamente 48% dos voluntários que executaram esta tarefa.

7.3.1 Cenas médias \bar{c}_m geradas através da Promediação Convencional

As cenas médias \bar{c}_m exibidas na Figura 7.9, são as cenas médias original e corrigida por um ser humano, geradas a partir do conjunto C de cenas semelhantes c_k sem filtragem de frequência.

A Figura 7.9 – A, mostra a cena média \bar{c}_m original, sem filtragem de frequência de uso. Nela, podemos observar que a cena gerada desrespeita totalmente as regras de posicionamento, sendo bastante tumultuada, com vários objetos sobrepostos, o que dificulta a visualização e entendimento da cena.

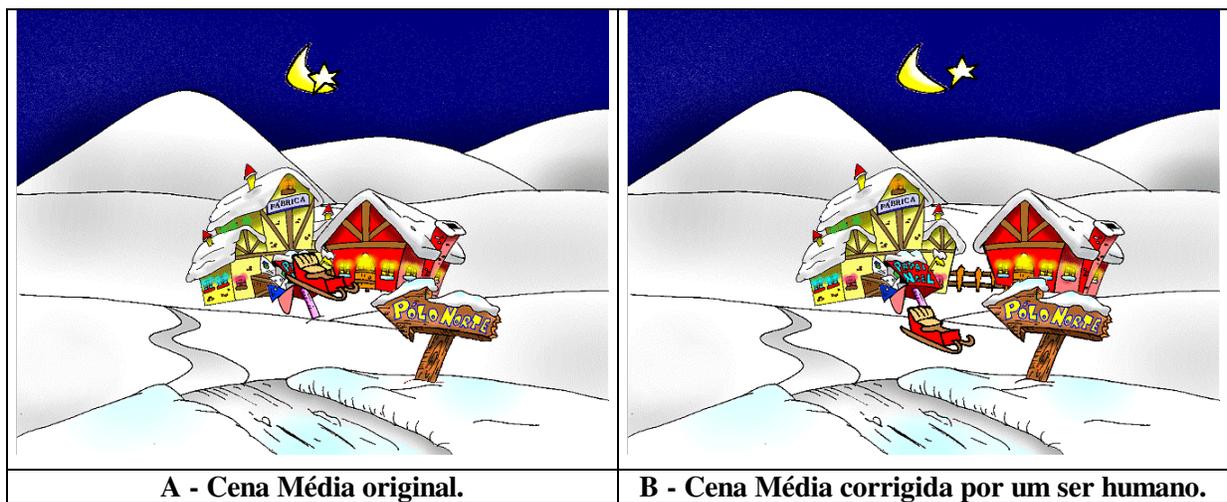


A - Cena Média original. **B - Cena Média corrigida por um ser humano Sem Filtros.**
Figura 7.9 - Cena Média original e corrigida por um ser humano Sem Filtros.

Para que a cena média \bar{c}_m mostrada na Figura 7.9 – A, possa ter uma coerência maior, facilitando o entendimento da cena, foram necessárias as seguintes correções (Figura 7.9 – B):

- Leve deslocamento do objeto “estrela” para a direita;
- Deslocamento do objeto “nuvem” para cima e para a esquerda;
- Deslocamento do objeto “trenó” para baixo;
- Deslocamento do objeto “arbusto” para a direita;
- Deslocamento do objeto “casa amarela” para a esquerda.

Aplicando o limiar de frequência de uso de objetos O_i superior a 11 indivíduos, ou seja, empregando os objetos O_i que tenham sido utilizados por mais de aproximadamente 48% dos usuários, eliminar-se-iam os objetos “conjunto de cartas”, “nuvem”, “carta voando” e “arbusto”, resultando numa cena mais limpa que não requer tantos ajustes, como pode ser observado na Figura 7.10 - A.



A - Cena Média original. **B - Cena Média corrigida por um ser humano com frequência de uso de objetos O_i superior a 11 indivíduos.**
Figura 7.10 - Cena Média \bar{c}_m original e corrigida por um ser humano com frequência de uso de objetos O_i superior a 11 indivíduos.

A Figura 7.10 – B mostra algumas correções feitas na Figura 7.10 – A, para que a cena ficasse com menor número de objetos sobrepostos, facilitando a visualização e aumentando a coerência. As correções feitas são descritas a seguir:

- O objeto “estrela” foi deslocado para a direita;
- O objeto “casa vermelha” foi deslocado para a esquerda;
- O objeto “trenó” foi deslocado para baixo;

7.3.2 Cenas prototípicas médias \bar{C} geradas através de Promediação Fuzzy

Na Figura 7.11 são mostrados todos os elementos pictóricos O_r disponíveis para serem usados como referência Z e suas quantidades de uso no conjunto C de cenas semelhantes c_k . Na Figura 7.11 foi desconsiderado o uso de filtros de frequência, visto que o objetivo é mostrar todos os elementos pictóricos que foram usados no conjunto C de cenas semelhantes c_k e as possíveis seleções da referência Z .



Figura 7.11 - Tela de seleção do elemento pictórico O_r para ser tomado como referência Z na tarefa “Fábrica de Brinquedos”.

Observou-se nesta tarefa que quanto mais longe estivesse a referência Z do centro da tela, pior o resultado da cena prototípica média \bar{c} gerada. Sendo assim, não foi possível a obtenção de boas cenas prototípicas médias \bar{c} quando foram selecionados elementos pictóricos O_r para a referência Z como por exemplo: “estrela”, “nuvem” ou “lua”, devido ao fato de sua localização média estar na região superior da tela.

Foi possível obter boas cenas prototípicas médias \bar{c} tomando como referência Z vários elementos pictóricos O_r e o centro da tela. Em alguns casos de cenas \bar{c} geradas, ocorreu de um objeto O_i desrespeitar as regras de posicionamento, ficando de alguma forma sobreposto à outro objeto. Porém, nesses casos pode-se retirar, se necessário, algum objeto O_i que esteja sobrepondo outro, desde que este não interfira no entendimento, nem na coerência da cena \bar{c} gerada.

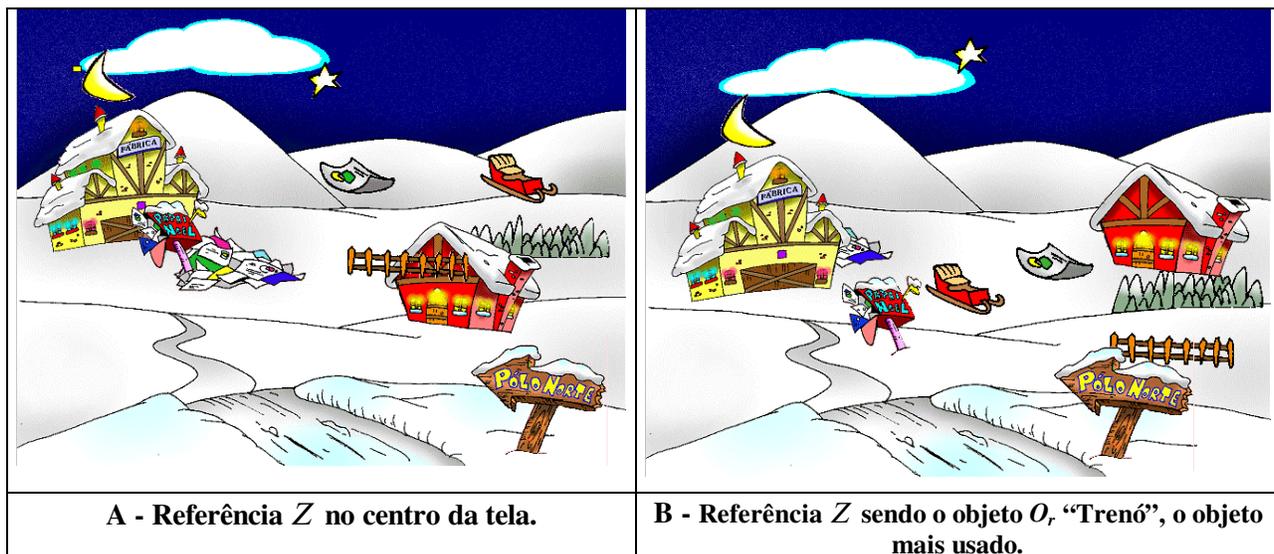


Figura 7.12 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Fábrica de Brinquedos” sem aplicação de filtragens.

A Figura 7.12 – A, mostra a cena prototípica média \bar{c} gerada, tendo como referência Z o centro da tela. Nela podemos observar com exceção do objeto O_i “cerca”, posicionado próximo ao telhado da casa vermelha, que todos os outros objetos O_i estão bem posicionados formando uma cena clara e coerente. Da mesma forma, na Figura 7.12 – B, cuja referência Z é o objeto O_i “trenó”, a cena também está bem montada.

A cerca que sobrepõe a casa na Figura 7.12 – A e a lua que sobrepõe a montanha na Figura 7.12– B, mostra um problema de ordem de colocação dos objetos. Se a casa estivesse à frente da cerca e a montanha à frente da lua, não haveria nenhuma infração às regras de posicionamento e a cena resultante seria melhor, como foi discutido no tópico 5.3.

A Figura 7.13 mostra a cena prototípica média \bar{c} gerada, tendo como referência Z , o centro da tela e o objeto O_i “trenó”, respectivamente, e onde foi aplicado um limiar de frequência de uso acima de 11 objetos usados.

Podemos observar através das Figura 7.12 e Figura 7.13, que o método da promediação fuzzy, permite a geração de cenas de melhor qualidade do que a promediação convencional (Figura 7.9 e Figura 7.10), visto que a primeira dispensa ajustes manuais, gerando cenas claras e bem montadas, que respeitam as regras de posicionamento.

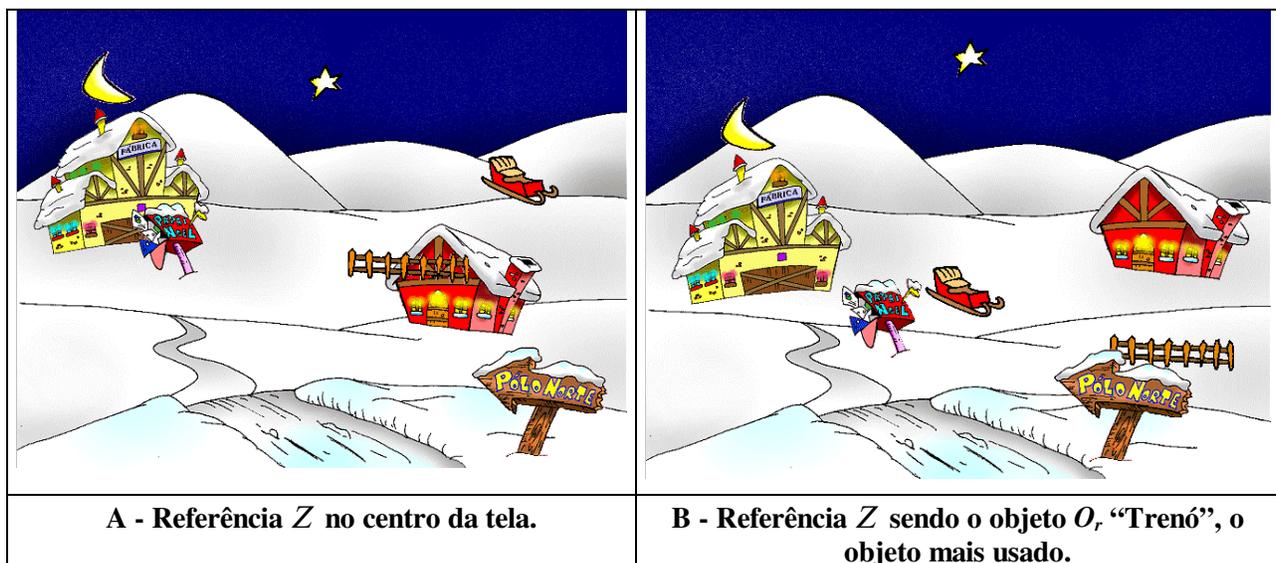


Figura 7.13 - Cenas prototípicas médias \bar{c} geradas sobre o tema “Fábrica de Brinquedos”, sendo o limiar de corte acima de 11 objetos.

7.3.3 Análise das Descrições l_k das Cenas c_k

Apenas dezoito dos vinte e três voluntários aceitaram fazer a descrição l_k do cenário montado. A Tabela 7.2 mostra a frequência com que os voluntários se referiram à casa do Papai Noel, às cartas que ele recebeu e à emoção dos sentimentos natalinos, como perdão, iluminação, felicidade, etc.

A moradia do Papai Noel foi descrita em 8 dos 18 textos criados pelos voluntários. As cartas e os pedidos de brinquedos foram referidos em 11 entre os 18 textos. Os sentimentos natalinos aparecem em 8 dos 18 textos. Entre os elementos pictóricos, o mais utilizado foi o objeto O_i “Trenó”, o qual não foi descrito em nenhum dos textos.

O objeto O_i “Caixa de correio” aparece 18 vezes, contra 11 vezes nos textos em que as cartas são mencionadas. Tanto o objeto O_i “Casa Vermelha”, quanto a “Casa Amarela” (Fábrica do Papai Noel)

apareceram 15 vezes na construção das cenas c_k , enquanto que nas descrições l_k , apenas 9 dos voluntários fizeram referência a elas.

CASA_MORA/[MORADIA DO PAPAÍ NOEL](-5,5,2,2)AAA& //PAPAÍ NOEL (-5,5,1)aaa&BBB& //NORTE (-5,5,1)bbb& \$ Frequência de ocorrências: 9 em 18 textos. 	
PEDIDOS_CARTA_PRES/[PRESENTES](-5,5,1,2)AAA& //PAPAÍ NOEL (-5,5,1)aaa&BBB& //Grande_Muit_Bastante (-5,5,1)bbb& \$ Frequência de ocorrências: 11 em 18 textos. 	
ACREDITA_BRILHO_ILUMI_PERDOA_FELI/[EMOCIONAL](-0,10,1,1)AAA& //NATAL_PAPAÍ NOEL (-5,5,1)aaa& \$ Frequência de ocorrências: 8 em 18 textos. 	
ELEMENTO PICTÓRICO	FREQÜÊNCIA
Nuvem	8
Trenó	19
Arbusto	10
Placa Pólo	17
Caixa Correio	18
Estrela	16
Lua	14
Cartas	11
Carta	8
Cerca	15
Casa Vermelha	15
Casa Amarela	15

Tabela 7.2 - Descrição da Cena Natalina.

7.4 Resumo do Capítulo

Neste capítulo, foram apresentados os resultados obtidos em cada atividade desenvolvida pelo grupo de voluntários. Estes resultados dizem respeito à geração das cenas médias \bar{c}_m e cenas prototípicas médias \bar{c} , assim como o resultado das análises das descrições l_k referente às cenas geradas pelo grupo de voluntários.

O próximo capítulo apresentará as conclusões tiradas desta pesquisa.

8 Conclusões

Pudemos observar no capítulo 7 - Resultados, que as cenas médias \bar{c}_m , geradas a partir do algoritmo de promediação convencional, não foram visualmente satisfatórias. Isso porque, não conseguimos obter cenas médias de qualidade que conseguissem exprimir uma idéia, principalmente porque, ao colocarmos cada objeto O_i em sua posição média, muitos objetos O_i acabam ficando sobrepostos, desrespeitando as regras de posicionamento e gerando desta forma uma cena confusa e carregada.

Portanto, a geração de cenas médias \bar{c}_m , a partir do cálculo de promediação pura e simples, não é uma metodologia eficiente, já que todas as cenas médias \bar{c}_m , geradas a partir deste artifício, tiveram de sofrer correções no posicionamento de vários objetos O_i para que se pudesse chegar a uma cena clara e coerente.

No entanto, quando utilizamos o algoritmo de promediação fuzzy, conseguimos obter cenas prototípicas médias \bar{c} visualmente satisfatórias e coerentes, ou seja, seus objetos obedeceram as regras de posicionamento.

O algoritmo de promediação fuzzy exige que seja definido um ponto de referência no espaço XY , o qual chamamos de referência Z . Esta referência é bastante flexível, podendo ser ou a posição central da tela, ou a posição média de um objeto O_i qualquer, presente num conjunto de cenas c_k . Isso acaba gerando duas grandes vantagens:

- Devido à flexibilidade da definição de um ponto de referência, podemos obter cenas prototípicas médias a partir de diferentes perspectivas, sem perder a coerência entre as cenas formadas;
- Mesmo não sendo possível a formação de cenas prototípicas médias \bar{c} de qualidade, usando todas as referências espaciais disponíveis (objetos O_i e o centro da tela), a probabilidade de se encontrar pelo menos uma referência que gere uma cena visualmente correta e coerente é grande, sendo que, em todos os testes feitos até o momento, foi possível gerar no mínimo uma cena média clara e bem formada, que não infringiu as regras de posicionamento.

Podemos observar através das cenas prototípicas médias \bar{c} , geradas a partir do algoritmo de promediação fuzzy, mostradas no capítulo 7, que obtivemos resultados satisfatórios, onde não foram necessárias correções nas cenas nem aplicação de filtros.

A partir das palavras mais utilizadas nas descrições I_k das cenas c_k , utilizando uma gramática G , pudemos identificar frases médias \bar{I} , que descrevem ou definem cenas prototípicas médias \bar{c} obtidas. Em outras palavras, obtivemos uma relação entre as frases médias \bar{I} e a composição da cena prototípica média \bar{c} , ambas a partir de um conjunto de cenas c_k e descrições destas cenas I_k , sobre um mesmo tema, para um grupo de voluntários k .

Portanto, do exposto acima, poderia-se concluir que:

1. A Promediação Fuzzy mostrou-se uma metodologia eficaz para a obtenção de uma cena prototípica média, já que com ela podemos obter uma cena final de qualidade, coerente e que respeite as regras de posicionamento;
2. Através da Promediação Fuzzy podemos gerar cenas prototípicas médias \bar{c} a partir de diferentes perspectivas, já que é necessário uma referência Z para a geração das cenas prototípicas médias \bar{c} e esta referência Z pode ser, tanto a localização do centro da tela como o ponto central da média de um objeto O_i qualquer, presente no texto;
3. Uma linguagem formal fuzzy $L(G)$ pode ser usada sobre uma gramática G para a análise de frases I_k que compõe;
4. Uma frase média \bar{I} pode ser uma descrição aceitável de uma cena prototípica média \bar{c} .

Com a TCP, Zadeh (Zadeh, 2001) propõe que é possível realizar uma grande variedade de tarefas físicas e mentais sem a necessidade de medições ou cálculos precisos. Os valores dos atributos são granulados, e cada grânulo reúne um grupo de valores (pontos, objetos) por similaridade, proximidade ou função, onde os grânulos de uma distância poderiam ser definidos como muito perto, perto, médio, longe, muito longe, dentre outros. Quando esses grânulos têm limites imprecisos são chamados F-Grânulos.

Portanto, a Teoria Computacional das Percepções (TCP) proposta por Zadeh (Zadeh, 2001), pressupõe que:

- As percepções humanas são organizadas a partir de F-Grânulos;
- São muitas vezes descritas em linguagem natural.

Na TCP, uma proposição p é vista como uma resposta a uma questão q e o significado de p representado como uma restrição generalizada R . O cálculo com percepções é proposto como um processamento de uma Linguagem de Restrições Generalizadas (LRG), de modo que a propagação das restrições tem como meta responder uma determinada questão (Zadeh, 2001).

Neste trabalho, a análise de uma cena visual utilizou questões q do tipo:

“Qual a localização de X em relação a Y ?”

e como solução à questão q obtivemos dois tipos de proposições:

- p_1 sobre as relações de distância entre os objetos que compõe a cena;
- p_2 sobre as direções entre esses mesmos objetos.

Em outras palavras, o reconhecimento visual de uma cena c_k envolveu o seguinte processamento em uma LRG entre objetos que compõe c_k :

$$c_k = [q_1, q_2, \dots, q_n], \text{ onde:}$$
$$q_i = (p_1, p_2) = \text{Localização}(O_x, O_y)$$

$p_1 =$ A distância entre (O_x, O_y) é {Muito Perto, Perto, Média, Longe, Muito Longe}

$p_2 =$ A direção entre (O_x, O_y) é {Norte, Sul, Leste, Oeste, Noroeste...}

De modo que a solução é encontrada com o tratamento das restrições:

$$R_1 = \{\text{Muito Perto, Perto, Média, Longe, Muito Longe}\}$$
$$R_2 = \{\text{Norte, Sul, Leste, Oeste, Noroeste...}\}$$

Dada uma família C de cenas semelhantes c_k , é possível calcular uma cena prototípica \bar{c} composta pelos objetos O_i mais freqüentes em C .

Desse modo, \bar{c} foi caracterizada pela disposição dos objetos O_i em relação a Z , porém ordenada por sua freqüência em C , isto é, os objetos mais freqüentes tiveram prioridade no seu posicionamento em \bar{c} e cada objeto O_i foi inserido nas soluções p_1 e p_2 de maior freqüência.

Portanto, o fundamento da Teoria Computacional das Percepções proposta por Zadeh (Zadeh, 2001), foi evidenciada neste trabalho, já que utilizando este conceito conseguimos obter cenas

prototípicas médias \bar{c} satisfatórias, que conseguem representar a idéia de um conjunto de cenas c_k , desenvolvidas por k indivíduos, sobre um determinado tema.

Através deste trabalho, foi possível também ampliar o sistema **SENSOR** (Serapião et al., 1996; Rocha, 1997).

Até então, o sistema **SENSOR** possuía os seguintes níveis de capacidade (Figura 8.1):

- Nível I: Extração de informações básicas sobre cores, luminância e contraste em uma imagem.
- Nível II: Reconhecimento dos atributos primários da imagem, tais como linhas, ângulos, descontinuidades de linhas, etc.
- Nível III: Reconhecimento de objetos a partir do conjunto de atributos reconhecidos.

Com a realização deste trabalho, foram aumentados mais dois níveis de capacidade no sistema **SENSOR** (Figura 8.1):

- Nível IV: Análise da relação de direção e distância entre os objetos O_i presentes numa cena c_k .
- Nível V: Geração de uma cena prototípica média a partir de um conjunto de cenas semelhantes c_k o que representa um **aprendizado**.

Vale ressaltar que o algoritmo de promediação fuzzy para geração de cena prototípica média, é um algoritmo de aprendizado, já que a partir de um conjunto de cenas semelhantes c_k consegue-se obter “cenas prototípicas” que, de acordo com o grau de restrição de frequência imposto, pode representar uma idéia que seja o consenso de um grupo de indivíduos k .

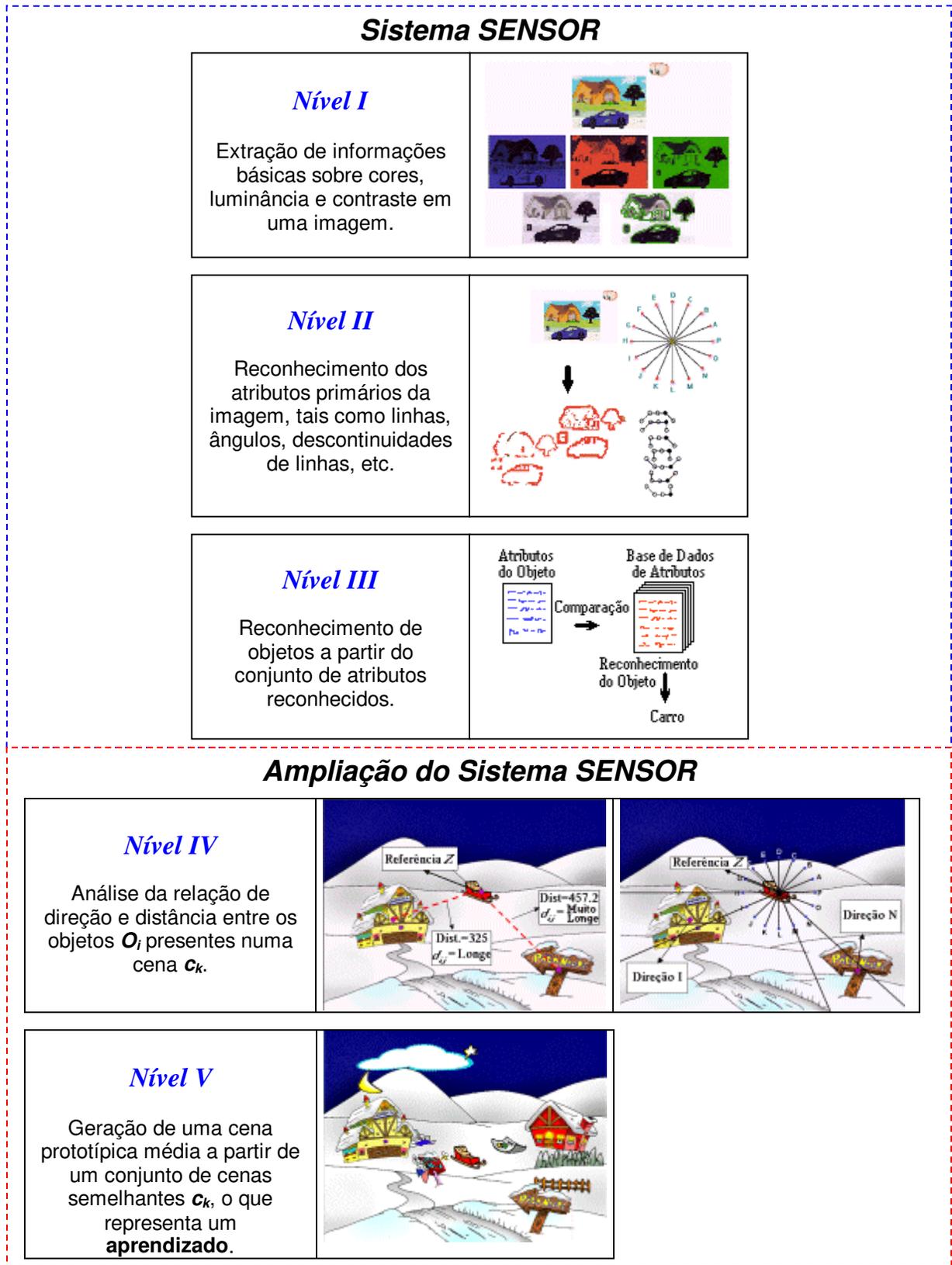


Figura 8.1 - Sistema SENSOR

Pode-se portanto, considerar satisfatório o resultado deste trabalho, uma vez que foi possível alcançar os objetivos propostos, como por exemplo:

1. Análise das relações de posicionamentos de elementos presentes em cenas \mathbf{c}_k ;
2. Análise das descrições I_k dessas cenas \mathbf{c}_k ;
3. Obtenção de cenas prototípicas médias \bar{c} e as correspondentes descrições \bar{I} de consenso;
4. Formalização dos agentes $\mathbf{o} \in \mathbf{O}$, através da TCP, capazes de aprender a identificar cenas prototípicas permitindo a expansão do sistema **SENSOR**.
5. Além disso, trouxe evidências de que a TCP (Teoria Computacional das Percepções), proposta por Lotfi A. Zadeh (Zadeh, 2001), pode ser utilizada na criação de modelos para a simulação do raciocínio humano.

8.1 Sugestões de Trabalhos Futuros

Gostaria de propor algumas sugestões de trabalhos futuros que dariam continuidade ao trabalho aqui apresentado :

- Utilizar outros métodos de criação de cenas prototípicas para efeito de comparação de desempenho com o algoritmo de promediação fuzzy apresentado.
- Ampliação do **VMC** para que ele seja capaz de reconhecer os objetos de uma cena utilizando o sistema **SENSOR**, possibilitando desta forma que qualquer tipo de cena seja analisada, independente do editor gráfico utilizado para gerá-las.
- Desenvolvimento de um algoritmo capaz de gerar automaticamente a referência Z para criação de uma cena prototípica \bar{c} com o menor número de infrações às regras possíveis.
- Ampliação do sistema para análise, classificação e geração de cenas prototípicas médias 3D.
- Explorar de forma mais abrangente a correlação visual - lingüística, ou seja, as relações entre uma cena visual e sua descrição escrita.
- Utilização do sistema **SENSOR** para criação de um sistema de busca de figuras num ambiente controlado, ou seja, a partir de uma base de dados com características de objetos (por exemplo: um garoto chamado Juca, uma menina chamada Laura, um robô, etc...) , seria possível encontrar imagens, numa base de figuras, passando como parâmetros os objetos que esta imagem deve conter e as características posicionais desejadas, como por exemplo:
 - ✓ Juca à direita de Laura;
 - ✓ Laura sozinha;
 - ✓ Juca e Laura perto do Robô.

9 Referências

Alegre, L., Rocha, A. F. and Morooka, C. K. "Intelligent approach to rod pumping problems". *SPE* 26253. 1993a, 249-255.

Alegre, L., Morooka, C. K. and Rocha, A. F. "Intelligent diagnosis of rod pumping problems". *SPE* 26516. 1993b, 97-107.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V.P., Karni, A., Prinster, A., Braun, A., Lalwani, A., Rauschecker, J.P., Turner, R., Neville, H. "Sentence Reading: A Functional fMRI Study at 4 Tesla". In: *Journal of Cognitive Neuroscience*. 1997; 9: 664-686.

Bhatnagar, S.C., Mandybur, G.T., Buckingham, H.W., Andy, O.J. "Language Representation in the Human Brain: Evidence from Cortical Mapping". In: *Language and Brain*. 2000; 74: 238-259.

Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E. "Making Memories: Brain activity that predicts how well visual experience will be remembered". In: *Science*. 1998; 281: 1158-1187.

Cabeza, R., Nyberg, L. "Imaging cognition: an empirical review of PET studies with normal subjects". In: *Journal Cognitive Neuroscience*. 1997; 9: 1-26.

Cabeza, R., Nyberg, L. "Imaging cognition II: an empirical review of 275 PET and fMRI studies". In: *Journal Cognitive Neuroscience*. 2000; 12: 1-47.

Chomsky, N. "Aspects of the Theory of Syntax". *MIT Press*. Cambridge, 1965.

Curtis, C.E., Zald, D.H., Pardo, J.V. "Organization of working memory within the human prefrontal cortex: a PET study of self-ordered object working memory". In: *Neuropsychology*. 2000; 38: 1503-1510.

Elias, L.J., Bulman-Fleming, M.B., McManus, I.C. "Visual temporal asymmetries are related to asymmetries in linguistic perception". In: *Neuropsychology*. 1999; 37: 1243-1249.

Gabrieli, J.D.E., Poldrack, R.A., Desmond, J.E. "The role of left prefrontal cortex in language and memory". In: *Proceedings of National Academy Science*. 1998; 95: 906-913.

Guilherme, I.R. "Sistema Jargão – Um sistema para Análise de Base de Dados em Linguagem Natural". *Tese de Doutorado, Universidade Estadual de Campinas – Faculdade de Engenharia Elétrica – Departamento de Computação e Automação*. 1996.

Holcomb, P.J., Kounios, J., Anderson, J.E., West, W.C. "Dual coding, context availability, and concreteness effect in sentence comprehension: An electrophysiological investigation". In: *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1999; 25: 741-742.

Holcomb, P.J., McPherson, W.B. "Event-related brain potentials reflect semantic priming in an object decision task". In: *Brain and Cognition*. 1994; 25: 259-276.

Hopcroft, J.E., and Ullman, J.D. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, MA, 1979.

Hopfinger, J.B., Buonocore, H.H., Magnun, G.R. "The neural mechanisms of top-down attentional control". In: *Nature Neuroscience*. 2000; 3: 284-291.

Hutsler, J.J., Gazzaniga, M.S. "The Organization of Human Language Cortex: Special Adaptation or Common Cortical Design?" In: *The Neuroscientist*. 1997; 3: 61-72.

Ishai, A., Ungerleider, L.G., Martin, A., Haxby, J.V. "The representation of objects in the human occipital and temporal cortex". In: *Journal of Cognitive Neuroscience* .12 suppl., 2000; 2: 35-51.

Johnsrude, I.S., Owen, A.M., Crane, J., Milner, B., Evans, A.C. "A cognitive activation study of memory for spatial relationships". In: *Neuropsychology*. 1999; 37: 829-841.

Kanwisher, N., Wojciulik, E. "Visual attention: Insights from brain imaging". In: *Nature Reviews: Neuroscience*. 2000; 1: 91-100.

Klir, G.J., Yuan, B. *Fuzzy Sets and Fuzzy Logic Theory and Applications*. Prentice Hall PTR, New Jersey, 1995.

Kosslyn, S.M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J.P., Thompson, W.L., Ganis, G., Sukeland, K.E., Alpert, N.M. "The role of area 17 in visual imagery: Convergent evidence from PET and rTMS". In: *Science*. 1999; 284: 167-170.

Maguire, E.A., Burgess, N., Donnett, J.G., Frackowiak, R.S.J., Fith, C.D., O'Keefe, J. "Knowing where and getting there: A human navigation network". In: *Science*. 1998; 280: 921-924.

Marshetz, C., Smith, E.E., Jonides, J., DeGutis, J., Chenevert, T.L. "Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms". In: *Journal of Cognitive Neuroscience*. 2000; 12 suppl, 2: 130-144.

Maunsell, J.H.R., Ferrera, V.P. "Attentional mechanisms in visual cortex". In: Gazzaniga, M.S. *The Cognitive Neuroscience*. Cambridge, The MIT Press, 1995, 451-461.

Mellet, E., Txourio-Mazoyer, N., Bricogne, S., Mazoyer, B., Kosslyn, S.M., Denis, M. "Functional anatomy of high-resolution visual mental imagery". In: *Journal of Cognitive Neuroscience*. 2000; 12: 98-109.

Meyer, M., Friederici, A.D., Von Cramon, D.Y. "Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands". In: *Cognitive Brain Research*. 2000; 9: 19-33.

Mizumoto, M., Toyoda, Y., Tanaka, k. "N-Fold Grammar". In: *Information Sciences*, 1973; 5: 25-43.

Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., Hatnao, K., Ito, K., Fukuda, H., Shcormann, T., Zilles, K. "Functional delineation of the human occipito-temporal areas related to face and scene processing". In: *Brain*. 2000; 123: 1903-1912.

Negoita, C.V., Ralescu, D.A. *Applications of Fuzzy Sets to Systems Analysis*. John Wiley & Sons, New York, 1975.

O'Craven, K.M., Kanwisher, N. "Mental imagery of faces and places activates corresponding stimulus-specific brain regions". In: *Journal of Cognitive Neuroscience*. 2000; 12: 1013-1023.

Paivio, A. "Dual coding theory: Retrospect and current status". In: *Canadian Journal of Psychology*. 1991; 45: 255-287.

Paivio, A. *Mental representations: A dual coding approach*. New York, Oxford University Press, 1996.

Perani, D., Cappa, S.F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M.M., Fazio, F. "The neural correlates of verb and noun processing: A PET study". In: *Brain*. 1999; 122: 2337-2344.

Pressman, R.S. *Engenharia de Software*. Makron Books, São Paulo, 1995.

Raj, T. "Patterns of brain during visual imagery of letters". In: *Journal of Cognitive Neuroscience*. 1999; 11: 282-299.

Riesenhuber, R., Poggio, T. "Hierarchical models of object recognition in cortex". In: *Nature Neuroscience*. 1999; 2: 1019-1025.

Rocha, A.F., Françoso, E., Balduino, M. A. "Neural Languages". In: *Fuzzy Sets and Systems*. 1980, 3/1, 11-35.

Rocha, A.F. "Neural Nets: A theory for brains and machines". In: *Lectures Notes in Artificial Intelligence*, Springer-Verlag, Berlin, 1992, vol. 638.

Rocha, A.F., Guilherme, I.R., Theoto, M., Miyadahira, A.M.K., Koizumi, M.S. "A Neural Net for Extracting Knowledge from Natural Language Data Bases". In: *IEEE Transactions on Neural Networks*, 1992, vol. 3, No. 5.

Rocha, A.F. "The brain as a symbol processing machine". In: *Progress in Neurobiology*, 1997;53:121-198.

Rocha, A.F., Rebello, M.P., Miura, K. "Toward a Theory of Molecular Computing". In: *Information Sciences*, 1998; 106: 123-157.

Rocha, A.F. *O cérebro: Um breve relato de sua função*. EINA, Jundiaí, 1999.

Rocha, F.T., Rocha, A.F., Rodella, E. "The Drawing Brain". In: *Journal of Cognitive Neuroscience*, Suppl: Cognitive Neuroscience Society Annual Meeting Program 2002, 2002; 145.

Sawrie, S.M., Martin, R.C., Gilliam, F.G., Faught, R.E., Maton, B., Hugg, J.W., Bush, N., Sinclair, K., Kuzniecky, R.I. "Visual confrontation naming and hippocampal function". In: *Brain*, 2000; 123: 770-780.

Serapião, A.B.S., Rocha, A.F., Rebelo, M.P. and Pedrycz, W. "Toward a theory of genetic systems". In: *Genetic Algorithms and Soft Computing*. Eds F. Herrera and J. L. Verdegay. Physica Verlag: Heidelberg, 1996; 68-94.

Searls, D.B. "The computational linguistics of biological scientists in Artificial Intelligence and Molecular Biology", In: *AAAI Press*, L. Henter, 1993, 47-120.

Smith, E.E., Jonides, J., Marshuetz, C., Koeppel, R. "Components of verbal working memory: Evidence from neuroimaging". In: *Proceedings of National Academy Science*. 1998; 95: 876-882.

St George, M., Kutas, M., Martines, A., Sereno, M.I. "Semantic integration in reading: engagement of the right hemisphere during discourse processing". In: *Brain*. 1999; 122: 1317-1325.

Ungerleider, L.G., Coutney, S.M., Haxby, J.V. "A neural system for human visual working memory". In: *Proceedings of National Academy Science*. 1998; 95: 883-890.

West, W.C., Holcomb, P.J. "Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation". In: *Journal of Cognitive Neuroscience*. 2000; 12: 1024-1037.

Young, M.P. "Open questions about the neural mechanisms of visual pattern recognition". In: Gazzaniga, M.S. *The Cognitive Neuroscience*. Cambridge, The MIT Press, 1995, 463-474.

Zadeh, L.A., "Fuzzy Sets". In: *Information and Control*. 1965; v 8(3): 338-353.

Zadeh, L.A. "The concept of a linguistic variable and its application to approximate reasoning I, II, III". In: *Information Sciences*. 1975, 8: 199-251, 301-357; 9: 43-80.

Zadeh, L.A. "The linguistic approach and its application to decision analysis". In: Ho, Y.C., Mitter, S.K.. *Directions in Large Scale Systems*. Plenum Press, New York, 1976, 339-370.

Zadeh, L.A. "A New Direction in AI – Toward a Computational Theory of Perceptions". In: *AI Magazine*. 2001, 73-84.

Zeki, S., Marini, L. "Three cortical stages of colour processing in the human brain". In: *Brain*. 1998; 121: 1669-1685.

Zimmermann, H.J. *Fuzzy Set Theory and Its Applications*. Kluwer Academic Publishers, 1996.

10 Índice Remissivo de Referências

Autor	Título do Trabalho	Página
Alegre et al., 1993a	Intelligent approach to rod pumping problems.	3, 24
Alegre et al., 1993b	Intelligent diagnosis of rod pumping problems.	3, 24
Bavelier et al., 1997	Sentence Reading: A Functional fMRI Study at 4 Tesla.	1
Bhatnagar et al., 2000	Language Representation in the Human Brain: Evidence from Cortical Mapping.	1
Brewer et al., 1998	Making Memories: Brain activity that predicts how well visual experience will be remembered.	2
Cabeza & Nyberg, 1997	Imaging cognition: an empirical review of PET studies with normal subjects.	1, 2
Cabeza & Nyberg, 2000	Imaging cognition II: an empirical review of 275 PET and fMRI studies.	1, 2
Chomsky, 1965	Aspects of the Theory of Syntax.	20, 22, 24
Curtis et al., 2000	Organization of working memory within the human prefrontal cortex: a PET study of self-ordered object working memory.	2
Elias et al., 1999	Visual temporal asymmetries are related to asymmetries in linguistic perception.	1
Gabrieli et al., 1998	The role of left prefrontal cortex in language and memory.	1
Guilherme, 1996	Sistema Jargão – Um sistema para Análise de Base de Dados em Linguagem Natural.	7, 35, 41, 45, 69
Holcomb et al., 1999	Dual coding, context availability, and concreteness effect in sentence comprehension: An electrophysiological investigation.	2
Holcomb & McPherson, 1994	Event-related brain potentials reflect semantic priming in an object decision task.	2
Hopcroft & Ullman, 1979	Introduction to Automata Theory, Languages, and Computation.	16, 17, 18, 19, 20, 22
Hopfinger et al., 2000	The neural mechanisms of top-down attentional control.	1
Hutsler & Gazzaniga, 1997	The Organization of Human Language Cortex: Special Adaptation or Common Cortical Design?	1

Ishai et al., 2000	The representation of objects in the human occipital and temporal cortex.	2
Johnsrude et al., 1999	A cognitive activation study of memory for spatial relationships.	2
Kanwisher & Wojciulik, 2000	Visual attention: Insights from brain imaging.	2
Klir & Yuan, 1995	Fuzzy Sets and Fuzzy Logic Theory and Applications.	12, 13, 14
Kosslyn et al., 1999	The role of area 17 in visual imagery: Convergent evidence from PET and rTMS.	2
Maguire et al., 1998	Knowing where and getting there: A human navigation network.	2
Marshetz et al., 2000	Order information in working memory: FMRI evidence for parietal and prefrontal mechanisms.	1
Maunsell & Ferrera, 1995	The Cognitive Neuroscience.	2
Mellet et al., 2000	Functional anatomy of thigh-resolution visual mental imagery.	2
Meyer et al., 2000	Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands.	1
Mizumoto et al., 1973	N-Fold Grammar.	22, 24
Nakamura et al., 2000	Functional delineation of the human occipito-temporal areas related to face and scene processing.	2
Negoita & Ralescu, 1975	Applicattions of Fuzzy Sets to Systems Analysis.	22, 24
O'Craven & Kanwisher, 2000	Mental imagery of faces and places activates corresponding sitmulus-specific brain regions.	2
Paivio, 1991	Dual coding theory: Retrospect and current status.	2
Paivio, 1996	Mental representations: A dual condign approach.	2
Perani et al., 1999	The neural correlates of verb and noun processing: A PET study.	2
Pressman, 1995	Engenharia de Software.	49, 65
Raij, 1999	Patterns of brain during visual imagery of letters.	2
Riesenhuber & Poggio, 1999	Hierarchical models of object recognition in cortex.	2
Rocha et al., 1980	Neural Languages.	24
Rocha, 1992	Neural Nets: A theory for brains and machines.	7, 22, 35, 41, 45

Rocha et al., 1992	A Neural Net for Extracting Knowledge from Natural Language Data Bases.	7, 35, 41, 69
Rocha, 1997	The brain as a symbol processing machine.	3, 18, 22, 23, 24, 29, 31, 94
Rocha et al., 1998	Toward a Theory of Molecular Computing.	18, 23
Rocha, 1999	O cérebro: Um breve relato de sua função.	3, 24, 27, 28
Rocha et al., 2002	The Drawing Brain.	2, 3, 37
Sawrie et al., 2000	Visual confrontation naming and hippocampal function.	2
Serapião et al., 1996	Toward a theory of genetic systems.	3, 24, 29, 31, 94
Searls, 1993	The computational linguistics of biological scientists in Artificial Intelligence and Molecular Biology.	24
Smith et al., 1998	Components of verbal working memory: Evidence from neuroimaging.	1
St George et al., 1999	Semantic integration in reading: engagement of the right hemisphere during discourse processing.	1
Ungerleider et al., 1998	A neural system for human visual working memory.	1, 2
West & Holcomb, 2000	Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation.	2
Young, 1995	Open questions about the neural mechanisms of visual pattern recognition.	2
Zadeh, 1965	Fuzzy Sets.	11, 14
Zadeh, 1975	The concept of a linguistic variable and its application to approximate reasoning I, II, III.	12
Zadeh, 1976	The linguistic approach and its application to decision analysis.	12
Zadeh, 2001	A New Direction in AI – Toward a Computational Theory of Perceptions.	5, 12, 13, 35, 45, 47, 92, 93, 96
Zeki & Marini, 1998	Three cortical stages of colour processing in the human brain.	2
Zimmermann, 1996	Fuzzy Set Theory and Its Applications.	12, 14