

Construção Automática de uma Hierarquia de Estratégias para Reconhecimento Tridimensional de Objetos Poliédricos

ANNA HELENA REALI COSTA RILLO

Divisão de Automação e Inteligência Artificial
Laboratório de Sistemas Integráveis (DAIA / LSI)
Escola Politécnica da Universidade de São Paulo
arillo@lsi.usp.br

Abstract. We developed a strategy-based computer vision system that can recognize three-dimensional objects from unknown viewpoints in single gray-scale images. The system is based on an off-line model preprocessing stage, where a strategy graph is automatically generated from 3D recognition-oriented object models. This strategy graph provides the representation of associations between features detected bottom-up and the data base of object models, enabling the on-line recognition algorithm to be particularly efficient by reducing the recognition to a 2D-matching process. This article focus on the definition and construction of the strategy graph.

1 Introdução

A despeito da aparente facilidade com que seres humanos são capazes de ver e entender o que está sendo visto, a visão tem se mostrado como sendo um processo extremamente complexo e ainda não completamente compreendido, mesmo no campo da fisiologia, neurologia, psicofísica e áreas afins, necessárias para o estudo da visão em seres biológicos [Overington (1992)]. Visão computacional procura entender e analisar este sentido.

Para contornar o problema imposto pela extrema complexidade da visão computacional, procurou-se identificar processos isolados, tratando-os como módulos independentes [Ballard-Brown (1982); Gonzalez-Wintz (1987)]. Desse modo, pode-se solucionar cada problema separadamente e então desenvolver metodologias de integração das informações fornecidas por estes processos independentes [Aloimonos (1989)].

Um dos processos fundamentais integrantes da visão computacional é o *Reconhecimento de Objetos*, pois envolve os problemas de *Identificação* e *Localização*, onde se procura definir quais objetos compõem determinada cena e qual o ponto espacial onde exatamente se encontram [Grimson (1990); Besl-Jain (1985); Lowe (1985)].

2 O Problema do Reconhecimento de Objetos

O mundo real que o ser humano vê e toca é composto primariamente por objetos. Se um objeto nunca visto antes for apresentado a um ser humano, ele é capaz de

extrair e organizar informações sobre o mesmo, procurando observá-lo sob diferentes pontos, visando conhecer sua geometria, textura, funcionalidade, etc.

Uma vez familiarizado com vários objetos, o ser humano normalmente pode identificá-los a partir de pontos arbitrários de observação, tendo não somente a capacidade de identificá-los, como também de localizá-los e qualitativamente descrever sua orientação.

Todo esse processo, de maneira ampla, pode ser descrito como o processo de Reconhecimento de Objetos. Analogamente, o objetivo do Reconhecimento de Objetos na Visão Computacional, doravante chamado simplesmente de Reconhecimento, consiste em, dados a informação sensorial e o conhecimento *a priori* sobre o mundo, identificar e localizar os objetos presentes na cena visualizada.

Muitos processos estão envolvidos na tarefa de Reconhecimento, a saber [Besl-Jain (1985)]:

- processo de formação da imagem;
- processo de descrição;
- processo de modelamento;
- processo de geração de hipóteses;
- processo de verificação.

O processo de formação da imagem implica em estabelecer que tipo(s) de sensor(es) é(são) usado(s) e quais princípios físicos estão envolvidos na criação da imagem.

O processo de descrição trata da definição de atributos relevantes da imagem e de seus respectivos algoritmos de extração. O trabalho aqui apresentado enfoca apenas *atributos geométricos*, chamados simplesmente primitivas ou atributos.

Essas primitivas são chamadas primitivas indexadoras, pois indexam a biblioteca de objetos, evidenciando somente objetos que as contenham em seus modelos. Dessa forma, seleciona-se objetos possuidores de determinadas primitivas, gerando hipóteses de correspondência entre primitivas da imagem e dos modelos, fornecendo, assim, objetos candidatos ao reconhecimento.

Essas hipóteses são confirmadas e refinadas, ou mesmo recusadas, no processo de verificação, gerando como respostas os objetos reconhecidos e suas respectivas localizações ou, então, realimentando o processo, requisitando novas buscas na imagem.

Esta abordagem, onde a descrição de uma imagem é comparada a entidades de uma biblioteca de modelos contendo a descrição, previamente computada, de cada objeto que pode fazer parte da cena, recebe o nome de Reconhecimento Baseado em Modelos.

Diversos aspectos são ressaltados como sendo importantes no Reconhecimento [Grimson (1990)]:

- tipo(s) de objeto(s);
- tipo de informação sensorial;
- representação da informação sobre cada objeto;
- armazenamento eficiente das representações em bibliotecas de objetos;
- métodos de aprendizado e de adição de novos objetos às bibliotecas;
- escolha de atributos (primitivas);
- métodos de extração de atributos das imagens;
- estratégia de correspondência adotada;
- métodos de indexação de objetos candidatos;
- processos para deduzir posição e orientação de objetos em imagens;
- estratégia usada para o processo de verificação;
- outros.

Nos últimos anos, um paradigma muito abordado tem sido o reconhecimento de objetos tridimensionais (3D) a partir de pontos de observação desconhecidos, utilizando simples imagens bidimensionais de intensidade luminosa (2D).

Reconhecimento 3D a partir de 2D envolve a projeção de uma cena em uma imagem, resultando num problema de difícil solução, o qual consiste na ambigüidade dos dados da imagem, devido ao fato de se perder a informação espacial durante o processo de projeção. Além desse fato, problemas inerentes a situações reais da visão computacional também são muito difíceis de se tratar, como oclusão, ruído, dados espúrios, etc. Apesar disso, a visão humana não apresenta dificuldades em propriamente inferir estruturas tridimensionais às imagens.

Buscando entender e simular este fenômeno, muitos sistemas de visão computacional têm sido propostos, os quais são capazes de identificar características de imagens

sob projeção e recuperar a posição e orientação de objetos a partir de uma única vista bidimensional. O trabalho aqui proposto é uma contribuição neste sentido, procurando dar uma nova e eficiente abordagem para o problema.

3 Trabalhos Correlatos e Proposta

Um significativo progresso tem sido realizado na área de reconhecimento de objetos 3D parcialmente oclusos em imagens 2D. Alguns exemplos importantes incluem: [Lowe (1985)], [Huttenlocher-Ullman (1988)], [Thompson-Mundy (1987)], [Lamdan-Wolfson (1988)], [Lamdan et al. (1988)], [Dickinson et al. (1992)], entre outros. Esses trabalhos diferem entre si basicamente nos seguintes tópicos:

- na complexidade dos atributos usados para indexação da base de dados (locais, intermediários, agrupamentos, primitivas volumétricas, etc)
- no método de indexação dos modelos (busca linear na base de dados, „hashing“ ou esquemas de votos, etc);
- no método de correspondência (por alinhamento, baseado em verificação, etc).

O sistema SCERPO de [Lowe (1985)] usa atributos invariantes com o ponto de observação para indexar a base de dados, num esquema de votos. Os votos são acumulados de acordo com atributos detectados na imagem, e usados para ordenar os modelos tridimensionais, com base no total de votos recebidos. As dificuldades apresentadas pelo sistema SCERPO são que os mecanismos de voto nem sempre são suficientemente discriminatórios, além de apresentar alto custo computacional durante transformações entre modelos 3D e atributos 2D no desenvolver do processo de correspondência.

[Huttenlocher-Ullman (1988)] propuseram a técnica de Alinhamento, onde uma hipótese para a transformação referente ao ponto de observação é realizada, baseada na correspondência de conjuntos mínimos de atributos e então todos os outros pontos do modelo são transformados e comparados com os pontos da imagem, tentando verificar a correspondência. Esta técnica apresenta a dificuldade de aplicar enumeração exaustiva sobre todos os possíveis pares de conjuntos mínimos do modelos e da imagem.

[Thompson-Mundy (1987)] efetuaram o reconhecimento através do agrupamento de todas as transformações („clustering“) derivadas de múltiplos pares de atributos. O reconhecimento então consiste na detecção de um ponto no espaço de possíveis transformações, cujas coordenadas mapeiem um número significativo de atributos da imagem às dos modelos. Essa técnica requer grande capacidade de armazenamento e é principalmente indicada para objetos poliédricos.

A técnica de "Hashing" geométrico, introduzida por Lamdan e outros [Lamdan-Wolfson (1988), Lamdan et al. (1988)], apresenta os mesmos problemas do sistema de Thompson e Mundy. Ela é baseada numa intensiva fase de processamento off-line, onde as informações dos modelos são armazenadas numa tabela ("hash-table"), usando características mínimas e invariantes com o ponto de observação. Isto permite um eficiente tempo de execução on-line, quando a base de dados não contém muitos objetos, pois, além de requerer uma enorme capacidade de armazenamento, efetua uma busca linear na tabela armazenada.

[Dickinson et al. (1992)] usaram aspectos para representar um vocabulário finito de primitivas volumétricas, a partir das quais os objetos podem ser construídos, deslocando a dificuldade maior do sistema para a detecção e agrupamento de primitivas volumétricas, as quais são muito mais difíceis de serem confiavelmente extraídas da imagem.

Acreditamos que a melhor solução consiste em um sistema híbrido, onde abordagens inerentemente dirigidas pelos modelos, usando particularidades dos mesmos para reduzir o processo de busca, são combinadas com outras, que utilizam-se de recursos mais genéricos, baseados nos dados extraídos da imagem para indexar os modelos. Assim, pode-se criar *estratégias* de reconhecimento mais eficientes do que as apresentadas até então. Estas estratégias devem utilizar *todo ou grande parte* do conhecimento disponível (características dos objetos e dos sensores usados, tarefa a ser executada, níveis desejados de soluções, etc), de modo *integrado*.

Para a condução da busca por correspondências, é necessário que a representação dessas estratégias seja de fácil acesso, permitindo diferentes níveis de abstrações para resolver o problema de reconhecimento de acordo com o detalhamento desejado. No entanto, as estratégias de reconhecimento devem ser indexadas por informações extraídas das imagens, de modo relativamente simples, restringindo o acesso a determinadas estratégias e selecionando outras, possibilitando uma redução do conjunto de estratégias a serem acessadas para a solução do problema.

Estes interesses foram aglutinados numa solução, resultando na proposta aqui apresentada: sistema RECTRI, onde a ênfase maior, neste trabalho é dada na construção de tal representação.

4 Sistema RECTRI

RECTRI integra um projeto em desenvolvimento na Divisão de Automação e Inteligência Artificial do Laboratório de Sistemas Integráveis da Escola Politécnica da USP (DAIA/LSI - EPUSP), chamado Célula Flexível de Montagem [Rillo et al. (1992)]. A

Célula de Montagem é formada por dois manipuladores, um sistema de planejamento de atividades e controle, um controlador lógico programável por GRAFCET, um sistema de visão binária [Rillo (1990)] e o sistema RECTRI [Rillo (1992)], o qual tem como função identificar e determinar a localização de peças tridimensionais em tarefas de montagem executadas pelos manipuladores.

O sistema RECTRI é composto por duas fases: pré-processamento e reconhecimento. Essa divisão é uma importante característica, uma vez que possibilita significativa redução no tempo de execução do algoritmo de reconhecimento por efetuar, em fase anterior, seleção e ordenação dos conhecimentos. A figura 1 ilustra as diversas etapas executadas nestas duas fases do sistema.

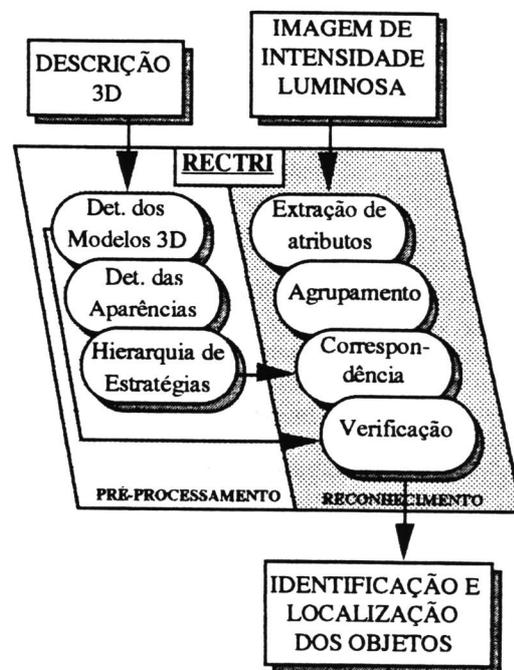


Figura 1 - Esquema geral do sistema RECTRI.

A fase de pré-processamento tem por função transformar as descrições 3D de cada objeto numa representação 3D interna, além de determinar o conjunto de vistas dos objetos e de definir uma estrutura para controle da busca por correspondências durante o processo de reconhecimento.

A fase de reconhecimento extrai atributos da imagem, agrupa-os, efetua a correspondência entre estes agrupamentos de atributos e aqueles dos modelos, sendo esta correspondência guiada pela estrutura definida na fase de pré-processamento. Finalmente, efetua a

verificação das hipóteses levantadas durante o processo de correspondência.

RECTRI utiliza-se extensivamente da fase de pré-processamento para buscar estratégias de reconhecimento. Estas estratégias estão representadas numa solução integrada, chamada *Hierarquia de Estratégias*, onde todos os objetos se encontram representados, com seus atributos ressaltados, evidenciando suas semelhanças e diferenças para conduzir eficientemente a busca por correspondências entre atributos da imagem e dos objetos. Essa hierarquia de estratégias constitui uma abordagem inovadora e promissora, sendo totalmente construída *de modo automático*.

Caracterizações qualitativas da morfologia das peças são refletidas nos atributos representados na hierarquia e naqueles extraídos das imagens. Esses atributos são agrupamentos escolhidos de primitivas, os quais atuam como indexadores da base de dados, selecionando estratégias eficientes para o reconhecimento.

O escopo deste trabalho se restringe à descrição e apresentação dos resultados referentes à fase de pré-processamento.

5 Fase de Pré-processamento

RECTRI utiliza um *modelo híbrido*, composto por três entidades básicas:

- *descrição tridimensional* de cada objeto (modelo do tipo volume-face-aresta-vértice);
- *representação multivista modificada* de cada descrição, dada pelo conjunto mínimo de *aspectos*;
- *Hierarquia de Estratégias*, onde todos os objetos se encontram representados e estratégias de busca por correspondência são definidas.

A fase de pré-processamento tem a função de construir esse modelo híbrido *automaticamente*, a partir de uma descrição 3D de cada objeto, editada pelo usuário. Esta descrição 3D dos objetos poderá também ser fornecida por um sistema CAD comercial, porém esta facilidade ainda não se encontra disponível na atual versão do sistema RECTRI.

A representação tridimensional dos objetos da base de dados utilizada neste trabalho consiste na representação *volume-face-aresta-vértice*, resultando em descrições usadas durante a *fase de verificação* do sistema e durante a *fase de construção do conjunto de vistas parciais* dos objetos. Esta descrição consiste em definir cada sólido por suas faces constituintes; cada face, por suas arestas constituintes e, cada aresta, por seus vértices e respectivas coordenadas.

6 Geração da Hierarquia de Estratégias

A geração da Hierarquia de Estratégias é realizada em três passos:

- definição dos conjuntos de aspectos,
- detecção dos atributos,
- construção da hierarquia.

6.1 Definição dos conjuntos de aspectos

O procedimento adotado para a definição das vistas parciais necessárias para o reconhecimento consiste em colocar cada objeto (modelo tridimensional) com seu centro de massa coincidindo com o centro de uma *esfera gaussiana de visualização*, a qual envolve o objeto.

Cada ponto na superfície da esfera gaussiana corresponde a um *ponto de observação* do objeto. A projeção do objeto no plano tangente a um ponto de observação da superfície da esfera de visualização corresponde a uma *vista* do objeto. Cada vista corresponde a uma relação ponto de observação-objeto.

Aglomerações de vistas isomórficas vizinhas resultam em *aspectos*, onde se estabelece uma relação região de observação-objeto. Os aspectos representam as projeções qualitativamente representativas, em relação à forma dos objetos visualizados e, dependendo da geometria do objeto, podem ser calculados analiticamente ou por amostragem.

Na definição dos aspectos por amostragem, um método muito difundido é o chamado método por simulação [Ikeuchi (1987); Dickinson et al. (1992)], utilizado neste trabalho e descrito adiante.

Todos os aspectos, de cada objeto, são utilizados em conjunto para a construção da Hierarquia de Estratégias.

6.2 Detecção dos atributos dos aspectos

Procedimentos de inferência de vários tipos são aplicados aos aspectos de todos os objetos, com a finalidade de extrair atributos úteis para o processo de reconhecimento.

Os atributos devem ser propriedades na imagem que podem ser formadas na ausência do domínio do conhecimento, porém precisam ser específicas o suficiente para servirem como termos de indexação. O principal critério para determinar a utilidade de um atributo é baseado nas seguintes qualidades:

- frequência com a qual é encontrado no objeto,
- robustez de sua detecção,
- complexidade de sua computação,
- característica discriminatória em relação aos outros atributos e
- probabilidade de corretamente identificar o objeto e sua localização.

Procedimentos idênticos aos aplicados à imagem, para a detecção de atributos na fase de reconhecimento,

são aqui utilizados. É importante salientar que os mesmos atributos 2D extraídos tanto da imagem quanto dos aspectos são utilizados no processo de correspondência e, portanto, armazenados na hierarquia de estratégias.

6.3 Construção da Hierarquia de Estratégias

O terceiro e último passo consiste em *organizar* os conjuntos de aspectos num grafo acíclico direcionado, chamado hierarquia, porém apresentando uma singularidade no sistema de herança, onde um sucessor poder ser apontado por mais de um antecessor (“tangled hierarquies”). Desse modo, os atributos comuns a diversos aspectos são evidenciados e representados somente uma vez na hierarquia, fornecendo um mecanismo simples e uniforme de *cooperação* entre *variações* e *similaridades* entre os objetos e entre os aspectos de um objeto.

A hierarquia é construída através de um *procedimento iterativo, incremental e totalmente automático*. O processo de construção parte das folhas da hierarquia até chegar à raiz, numa direção *inversa* à utilizada durante o reconhecimento.

Cada folha está associada a um aspecto único de um objeto ou a aspectos cujas diferenças em suas projeções não são representáveis e/ou perceptíveis ao sistema.

Cada folha representa *conectividade* (ou *inclusão*) de atributos extraídos e selecionados dos aspectos. *Iterativamente* retira-se um atributo representado no nó analisado, formando outros nós (portanto, com um menor número de atributos), até alcançar os *nós mínimos*, formados por um só atributo.

Um critério de *validade* é aplicado a cada nó criado. Considera-se como *válidos* somente os nós que possuam *conectividade* entre todos os seus atributos. Este fator é importante no controle da busca por atributos conectados na imagem, durante a fase de reconhecimento, e está baseado na suposição de que atributos conectados são mais prováveis de pertencerem a um mesmo objeto.

Cada nó válido criado é testado com todos os outros de mesmo número de propriedades já existentes na hierarquia. Nós equivalentes são *aglomerados* em um só nó. A equivalência de dois nós é determinada por meio de uma busca recursiva numa *árvore de equivalência* entre atributos pertencentes aos nós testados. Para cada par de nós testados é construída, automaticamente, uma *árvore de equivalência*. Cada par de atributos equivalentes determinados pela árvore são testados, segundo três critérios:

- análise de *conectividade* entre os atributos (atributos vizinhos devem ser do mesmo tipo e a conectividade também de mesmo tipo e quantidade),

- análise da *condição de contorno* entre ambos os nós, para cada atributo agregado,
- análise do *paralelismo* mantido pelos segmentos dos atributos pertencentes aos nós.

Durante todo este processo de construção, diversos critérios são utilizados (alguns implicitamente, outros, de modo explícito) para organizar a hierarquia, considerando o mecanismo utilizado de controle da busca (no caso, foi usado a *busca em profundidade*), objetos da base de dados, características da tarefa, etc, enfim, todas as informações disponíveis. No sistema RECTRI a organização dos nós da hierarquia é feita em função do controle da busca e da utilidade dos atributos constituintes de cada nó.

A *organização* da hierarquia, juntamente com uma *terminação heurística* da busca por correspondência durante a fase de reconhecimento, visa uma diminuição do espaço de busca, resultando num menor tempo de reconhecimento.

Deve-se ressaltar que um processo de aprendizado pode interagir nessa organização, visando maior eficiência, conforme adquira experiência. Deve-se notar também que, apesar do processo de construção da hierarquia ser computacionalmente custoso, ele é executado no pré-processamento, portanto, uma única vez, para cada banco de dados, permitindo uma maior eficiência no reconhecimento.

7 Resultados

O domínio escolhido para a implementação atual do RECTRI foi para reconhecimento dos objetos poliédricos apresentados na figura 2.

Objetos poliédricos foram escolhidos pelo fato de que sistemas que os reconhecem possuem a maioria das características essenciais para análise e interpretação de uma cena, como formação da imagem, modelos, relações espaciais entre os objetos da cena (oclusão, sobreposição, etc), mas, no entanto, são bastante simples para modelar tridimensionalmente, extrair atributos, efetuar correspondência espacial, etc.

As primitivas usadas na versão atual do sistema RECTRI são segmentos de reta, as quais são agrupadas em duas etapas. Primeiramente, primitivas que satisfaçam critérios de colinearidade, paralelismo e junção entre si são agrupadas. Destes grupos são construídos outros grupos, mais complexos, procurando outras relações entre os primeiros; assim, são criados grupos de polígonos com oito arestas, paralelogramos, trapézios, triângulos, cinco segmentos conectados em cadeia aberta, quatro segmentos conectados em cadeia aberta, entre outros. Esses agrupamentos são classificados de acordo com os critérios citados no item 6.2.

Hierarquias de estratégias são geradas para o reconhecimento dos objetos da figura 2. Para efeito de visualização de uma hierarquia gerada, a figura 3

apresenta a hierarquia criada para os aspectos determinados para as peças L e Prisma, considerando 10 pontos sobre a esfera de visualização.

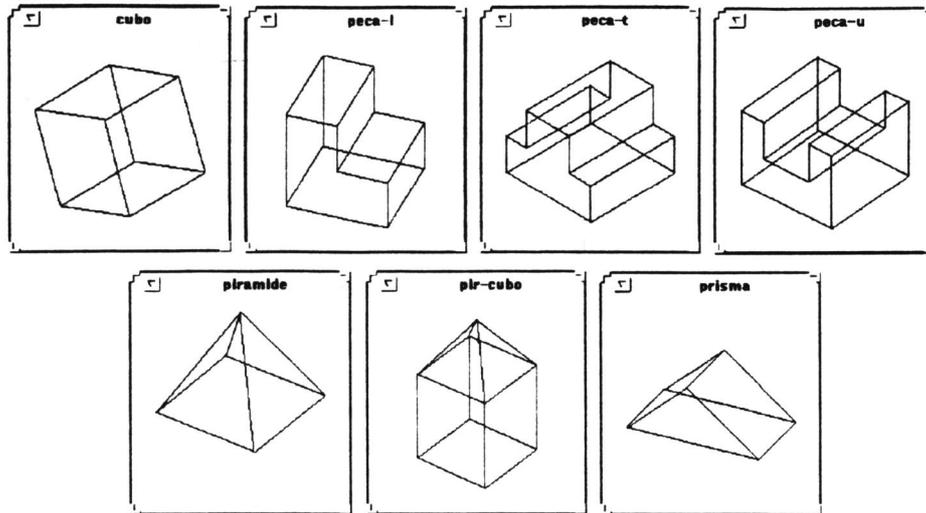


Figura 2 - Conjunto de peças.

8 Conclusões

Neste trabalho foi proposto um modelo computacional para o reconhecimento de objetos rígidos tridimensionais, parcialmente oclusos, a partir de uma única imagem bidimensional de intensidade luminosa refletida por uma cena, modelo este baseado na seleção e organização de características úteis e discriminatórias. Este modelo foi implementado num sistema de reconhecimento de peças poliédricas, chamado RECTRI.

A descrição tridimensional (volume-face-aresta-vértice) de cada peça se torna a única fonte de informação para o sistema RECTRI. Nesta implementação, somente informações geométricas e topológicas são utilizadas.

Na construção do hierarquia de estratégias, características muito pequenas, descrevendo detalhes dos objetos, são automaticamente eliminadas ou recebem prioridade menor, sendo somente utilizadas no processo de verificação. Entretanto, se objetos similares coexistem em uma cena, uma eficiente estratégia é construída para discriminar esses casos.

Nossos atuais esforços consistem em um desenvolvimento mais profundo do sistema de reconhecimento. Há muito espaço ainda para pesquisas voltadas para segmentação da imagem, com ênfase em algoritmos mais robustos e eficientes.

Outra importante área de investigação consiste no aprendizado a partir de cenas corretamente interpretadas, uma vez que a hierarquia gerada pode ser continuamente

reorganizada e todos os parâmetros utilizados no sistema RECTRI podem ser ajustados, procurando uma classificação ótima e uma localização mais eficiente, com o ganho de experiência do processo visual.

Nossa proposta ofereceu uma possível solução para o problema de reconhecimento a partir de modelos, mas acreditamos que outros tipos de representação devam ser exploradas, tanto para as aplicações às quais nosso sistema está voltado quanto para outras mais gerais, uma vez que o uso de sofisticadas estruturas de dados é um importante componente de sistemas de visão computacional.

8 Referências Bibliográficas

- Aloimonos, J.Y. Unification and integration of visual modules: an extension of the Marr paradigm. In: DARPA Image Understanding Workshop, Palo Alto, California, May 23-26, 1989. Proceedings. DARPA, 1989. p.507-51.
- Ballard, D.H.; Brown, C.M. Computer vision. Englewood Cliffs, Prentice-Hall, Inc., 1982.
- Besl, P.J.; Jain, R.C. Three-Dimensional Object Recognition, ACM Computing Surveys, Vol. 17, No. 1, pp. 75-145, 1985.
- Dickinson, S.J.; Pentland, A. P.; Rosenfeld, A. 3D Shape Recovery Using Distributed Aspect Matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, IEEE, pp. 174-198, 1992.

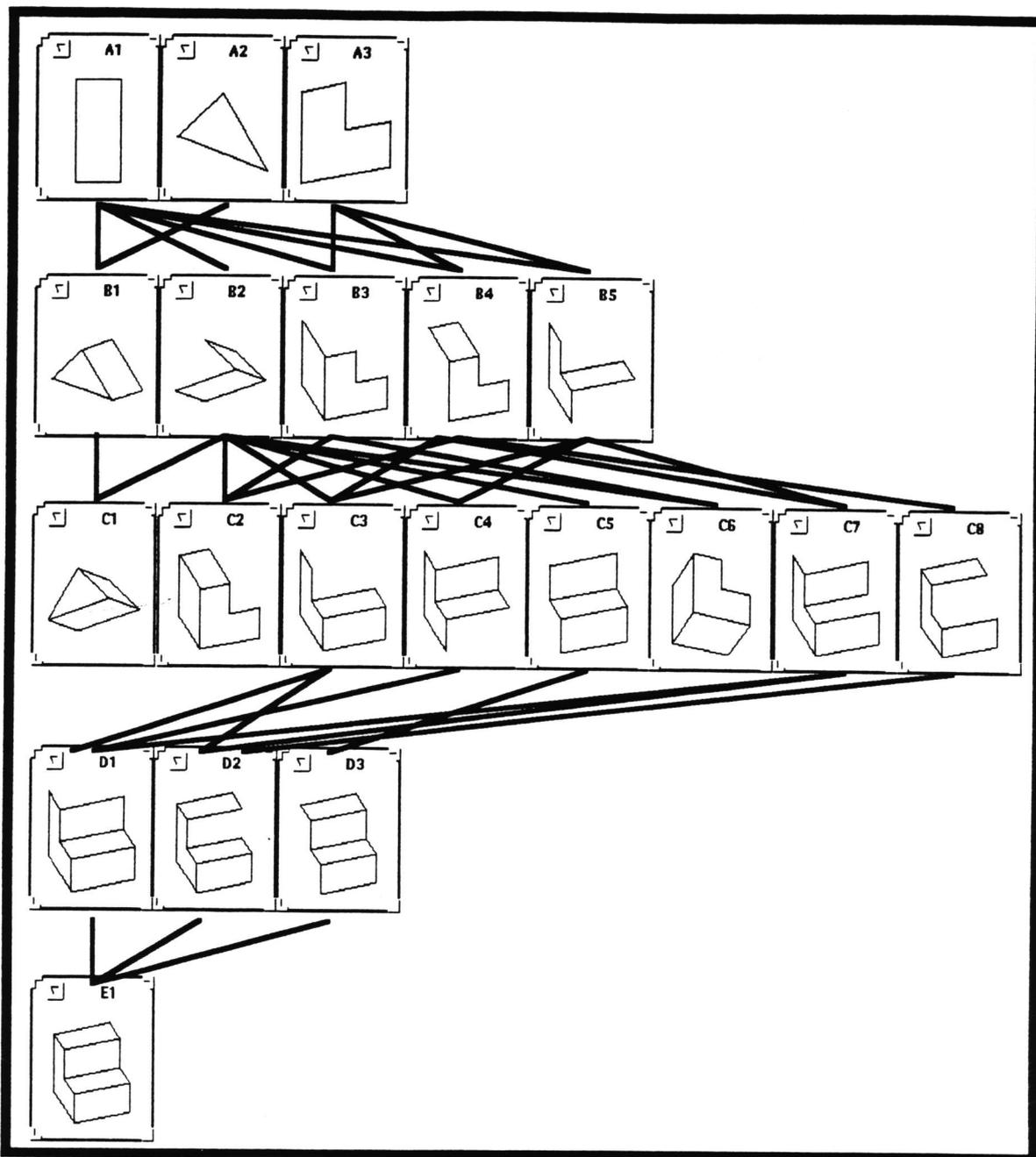


Figura 3 - Hierarquia de estratégias de reconhecimento gerada automaticamente a partir dos aspectos determinados para as peças L e Prisma da figura anterior.

- Gonzalez, R.C.; Wintz, P. Digital image processing. Reading, Addison-Wesley Publishing Company, 1987.
- Grimson, W.E.L. Object Recognition by Computer: the Role of Geometric Constraints. Cambridge, The MIT Press, 1990.
- Huttenlocher, T.; Ullman, S. Recognizing Solid Objects by Alignment, Proc. of the DARPA Image Understanding Workshop, Cambridge, Massachusetts, pp. 1114-1122, April 1988.
- Ikeuchi, K. Model-Based Interpretation of Range Imagery, in DARPA Image Understanding Workshop (Proceedings), pp. 321-339, 1987.
- Lamdan, Y.; Wolfson, H.J. Geometric Hashing: A General and Efficient Model-based Recognition Scheme, in International Conference on Computer Vision (Proceedings), Tampa, Florida, pp.238-249, 1988.
- Lamdan, Y.; Schwartz, J. T.; Wolfson, H. J. On Recognition of 3D-Objects from 2D-Images, in IEEE International Conference on Robotics and Automation (Proceedings), Philadelphia, IEEE, pp. 1407-1413, 1988.
- Lowe, D.G. Perceptual Organization and Visual Recognition, Kluwer Academic Publishers, Boston, MA, 1985.
- Overington, I. Computer vision: a unified, biologically-inspired approach. Amsterdam, Elsevier, 1992.
- Rillo, A.H.R.C. Um sistema de visão binária para reconhecimento de peças isoladas e parcialmente oclusas. III Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens (III SIBGRAPI). Anais. Gramado, RS, 1990, pp. 236-245.
- Rillo, A.H.R.C. Grouping-based recognition system, in Model-Based Vision: Development and Tools (Proceedings), R. M. Larson & H. N. Nasr (eds.), Proc. SPIE 1609, 1992, pp. 274-282.
- Rillo, M.; Rillo, A.H.R.C.; Costa, L.A.R. The LSI Assembly Cell, in 7th IFAC/IFIP/IFORS/IMACS/ISPE Symposium on Information Control Problems in Manufacturing Technology (Proceedings), May 25-28, 1992, Toronto, Canada, 1992.
- Thompson, D.W.; Mundy, J.L. Three-Dimensional Model Matching from an Unconstrained Viewpoint", in IEEE International Conference on Robotics and Automation (Proceedings), Raleigh, NC, IEEE, pp. 208-220, 1987.