

UM SISTEMA DE VISÃO BINÁRIA PARA RECONHECIMENTO DE PEÇAS ISOLADAS E PARCIALMENTE OCLUSAS

Anna Helena Reali Costa Rillo
Laboratório de Sistemas Integráveis
Depto. de Eng. de Eletricidade, EPUSP
Av. Prof. Luciano Gualberto, 158, Trav.3
05508 São Paulo, SP

RESUMO - Este trabalho descreve um sistema de visão artificial para robôs industriais utilizados em tarefas de montagem, o qual permite identificar e localizar objetos complexos numa cena contendo vários objetos, isolados ou com oclusão parcial dos mesmos, utilizando a imagem binária da cena.

1. INTRODUÇÃO

A utilização de robôs em sistemas automáticos de manufatura está fazendo com que uma especial atenção seja dedicada a sistemas de visão artificial.

A maioria dos sistemas de reconhecimento de peças industriais são sistemas baseados em modelos ("model-based systems"), nos quais o reconhecimento se dá através do casamento da imagem de entrada com um conjunto de modelos pré-definidos das peças, determinando qual é a peça, sua posição e orientação em relação a um referencial. Um crescente número de estudos tem sido conduzido, investigando vários métodos para reconhecimento de peças industriais. Este trabalho vem dar mais uma contribuição neste sentido.

O sistema implementado [1] é a base de um sistema de visão que faz parte de uma célula flexível de montagem que está sendo desenvolvida na Divisão de Automação do Laboratório de Sistemas Integráveis do Departamento de Engenharia de Eletricidade da Escola Politécnica da Universidade de São Paulo.

2. DESCRIÇÃO GERAL

No sistema proposto, o sistema de visão está acoplado a um robô de montagem, fornecendo a ele informações referentes à identidade e localização das peças, para que este possa manipulá-las e realizar a tarefa de montagem.

As peças necessárias à montagem, as quais são colocadas arbitrariamente sobre uma mesa, são apresentadas ao sistema de visão, o qual capta a imagem através do Módulo de Aquisição de Imagens Coloridas [2], interpreta e determina as informações necessárias ao reconhecimento.

O número de peças que podem aparecer na cena, assim como seus respectivos modelos, são previamente conhecidos e determinados.

A maioria dos sistemas de visão industrial disponíveis pode somente localizar e reconhecer peças isoladas. Entretanto, numa cena industrial típica, as peças podem estar em contato ou parcialmente oclusas, de modo que o fato de considerar somente objetos isolados pode resultar numa limitação muito forte. A fim de evitar esta restrição, este sistema de visão artificial foi desenvolvido para permitir não apenas o reconhecimento de peças isoladas na cena, mas também de peças que se tocam ou se superpoem, desde que partes consideráveis delas ainda estejam visíveis.

Devido ao fato do sistema de visão proposto detectar somente silhuetas (sistema de visão binária bidimensional), os objetos têm que possuir silhuetas que permitam uma única determinação de orientação ou que sejam rotativamente invariantes. Além disso, esta silhueta tem que ser suficiente para que a peça seja identificada.

3. DESCRIÇÃO FUNCIONAL

O sistema de visão artificial proposto é composto por duas fases distintas: FASE DE TREINAMENTO e FASE DE EXECUÇÃO.

O principal interesse na definição de uma fase de treinamento é facilitar e agilizar ao máximo a fase de execução do sistema. Esta fase de treinamento só é processada na inicialização do sistema, quando os objetos são apresentados e definidos para o sistema (aprendizado), ou ainda na reprogramação do mesmo, quando os objetos são alterados ou novos objetos introduzidos. Suas funções são descritas no decorrer do trabalho, juntamente com a descrição dos passos da fase de execução.

A fase de execução é processada quando o sistema todo está em regime. Ela é composta pelos seguintes passos: aquisição da imagem, binarização, extração dos contornos, segmentação, reconhecimento das peças isoladas e reconhecimento das peças em contato.

PASSO 1 - Aquisição da Imagem:

A cena é captada por uma câmera de vídeo e sua imagem é digitalizada em níveis de cinza pelo Módulo de Aquisição de Imagens Coloridas [2].

PASSO 2 - Binarização da Imagem:

A partir da imagem digitalizada em níveis de cinza, é gerada uma imagem binária, a qual possui somente dois níveis de luminosidade: preto ou branco.

A binarização é feita comparando-se o valor de luminosidade de cada pixel com um valor limite. A definição deste valor limite é feita na fase de treinamento do sistema, através da análise de um histograma dos níveis de cinza da imagem.

A vantagem deste tipo de imagem é a grande facilidade de processamento que a mesma oferece, permitindo a separação fácil do objeto de seu fundo. No entanto, não se deve esquecer suas limitações, entre elas: perda de informação ao transformar uma imagem em binária, necessidade de alto contraste entre objeto e fundo e tratamento essencialmente bidimensional dos objetos.

PASSO 3 - Extração do contorno:

De posse da imagem binária da cena, o próximo passo da fase de execução do sistema de visão será extrair o contorno das peças da imagem.

A extração dos contornos das peças na matriz de imagem é feita considerando, como unidade básica para análise, quatro pixels adjacentes, caminhando pixel a pixel (passo de um pixel), tanto nas colunas quanto nas linhas da matriz [3]. Cada pixel será rebatizado como sendo um ponto, englobando as características dos quatro pixels da unidade básica para análise.

As seis possíveis unidades básicas para análise dos elementos da linha de contorno são as apresentadas na figura 1.

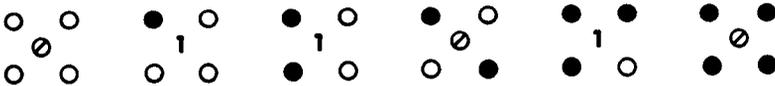


Figura 1 - Unidades básicas para definição do contorno.

PASSO 4 - Segmentação da Imagem

Este algoritmo consiste em determinar, a partir dos contornos, quais os pontos que compõem determinado objeto na imagem para que seus atributos possam ser determinados.

O raciocínio básico é que um ponto na matriz da imagem é definido como sendo conectado a outro ponto se ele ocupar uma das oito locações imediatamente adjacentes (figura 2).

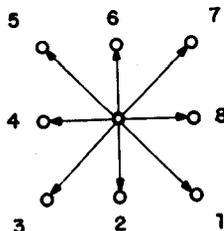


Figura 2 - Cadeia de vetores elementares.

Portanto, a passagem de um ponto a outro conectado deve ser feita por um dos oito vetores elementares assim definidos.

O algoritmo só considera contornos fechados. Por isso, quando alguns pontos forem encontrados mas não resultarem num contorno fechado, eles serão ignorados.

PASSO 5 - Reconhecimento das peças isoladas

a) Identificação das peças isoladas na cena:

Com a obtenção da imagem segmentada, o sistema extrai atributos que possibilitarão a identificação das peças isoladas na cena.

A classificação do objeto é feita baseada nestes atributos extraídos, que são: perímetro total, área total, área total dos buracos na peça, perímetros e áreas do maior e do menor buracos da peça.

Através de uma amostra de treinamento constituída de objetos cuja identidade é conhecida, determinam-se valores típicos para os atributos considerados. A classificação é feita comparando-se os valores dos atributos dos objetos cuja identidade se quer determinar com valores armazenados ou então com estatísticas destes valores, definidos na fase de treinamento.

Estas comparações são determinadas por uma árvore de decisão binária, construída automaticamente na fase de treinamento do sistema. Após percorrer toda a árvore, ter-se-á a identificação das peças isoladas na cena, bem como o número e modelos correspondentes às peças que não puderam ser reconhecidas nesta primeira fase do processo por estarem em contato entre si.

b) Determinação da Posição das Peças Isoladas:

Uma vez identificadas as peças isoladas da cena, a posição destas é determinada pelas coordenadas absolutas x e y do centro de área das mesmas (centróide).

O centro de área é o centro de massa de uma figura de mesmo contorno, com massa constante por unidade de área. Por sua vez, o centro de massa é aquele ponto onde toda a massa do objeto poderia ser concentrada, sem mudar o momento de primeira ordem do objeto em relação a qualquer eixo. Portanto, determinamos as coordenadas absolutas x e y do centro de área do objeto a partir dos momentos de primeira ordem das peças (ou momentos estáticos da área).

c) Determinação da Orientação das Peças Isoladas:

A orientação de um objeto é que define como o objeto se encontra no nosso campo de visão.

Se considerarmos objetos que apresentem um certo grau de alongamento, pode-se definir a orientação do objeto como sendo a orientação do eixo de alongamento. A prática usual é escolher o eixo do menor momento de segunda ordem como sendo o eixo de alongamento [4]. No plano bidimensional, isto é equivalente ao eixo de menor inércia.

Deve-se considerar, entretanto, que o eixo de menor inércia simplesmente define uma linha que passa através do objeto, não determinando um sentido à ela. Definiu-se, então, um ponto para determinar o sentido do eixo de menor inércia da peça. O ponto é aquele que se encontra mais distante do centróide, sobre o eixo de menor inércia.

Na fase de treinamento do sistema é calculada a distância do centróide ao ponto P mais distante sobre o eixo de menor momento de inércia, a fim de facilitar os cálculos para a determinação de P na fase de execução do sistema.

PASSO B - Reconhecimento das peças em contato:

Neste passo do processo, as peças que se encontravam isoladas na cena já foram reconhecidas, restando somente na imagem peças que se encontram em contato entre si.

O número de peças a serem identificadas neste passo é conhecido, dado que se tem posse do número total de peças na cena (fornecido pelo usuário na fase de treinamento) e do número de peças isoladas na cena (determinado no passo anterior da fase de execução).

O método utilizado é baseado no exposto em [5,6] e consiste em comparar segmentos de tamanho fixo do contorno do modelo de um objeto (o tamanho é estipulado em pixels pelo usuário na fase de treinamento), com segmentos de mesmo tamanho do contorno da imagem.

Os segmentos a serem comparados são segmentos selecionados do contorno dos modelos das peças, de tal modo que possam inequivocadamente definir um objeto, sua posição e orientação, dados os outros objetos que podem ocorrer na imagem. Estes segmentos são ditos segmentos salientes e são determinados automaticamente na fase de treinamento do sistema.

O procedimento de reconhecimento busca por um objeto na imagem contendo objetos parcialmente oclusos primeiramente tentando localizar o mais saliente dos segmentos daquele objeto.

Se este segmento for encontrado, a identidade, a posição e orientação do objeto em questão ficam definidas, uma vez que a fase de treinamento seleciona somente aqueles segmentos que definam um único objeto e seu posicionamento. Caso contrário, o procedimento continua com o próximo segmento mais saliente daquele objeto procurado.

Se nenhum dos segmentos salientes selecionados for encontrado, é enviada uma mensagem: "Reconhecimento impossibilitado: alto grau de oclusão".

Para se efetuar a comparação entre os segmentos, os contornos do modelo e da imagem são representados em dois espaços: no espaço cartesiano ($x - y$) e no espaço ângulo-comprimento ($\theta - s$).

A transformação para o espaço θ - s é feita a partir da cadeia de vetores elementares dos contornos, pixel a pixel, sendo que o ângulo é definido como sendo o valor médio dos ângulos dados pelos dois vetores elementares que atuam no pixel.

Durante a comparação, a representação θ - s de um segmento saliente do modelo é transportada pelo eixo s , alinhando seu centro ao centro do segmento da imagem a ser comparado. O segmento do modelo é então transportado na direção θ , até que o ângulo médio do segmento do modelo e da imagem tenham o mesmo valor θ . A diferença entre o θ médio inicial e o θ médio final do segmento do modelo é chamada "ângulo de casamento".

A diferença em θ é encontrada a partir da correspondência feita entre cada ponto dos segmentos da imagem e do modelo. O inverso da soma dos quadrados destas diferenças é usado para medir a similaridade entre os dois segmentos.

Se eles forem similares, considerando uma faixa de variação pré-estabelecida pelo usuário na fase de treinamento do sistema, assume-se que o segmento saliente do modelo foi encontrado na imagem. Caso contrário, desloca-se o segmento saliente procurado de "ne" pixels no eixo s e prossegue-se analogamente a busca, sendo "ne" o número de pixels para espaçamento de busca, fornecido pelo usuário na fase de treinamento do sistema (fig.3).

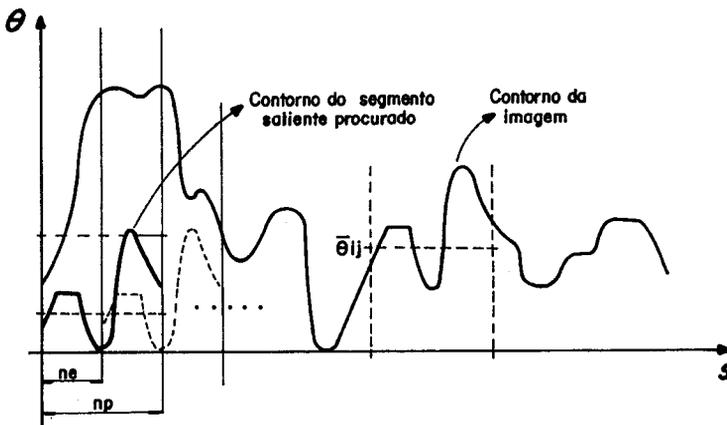


Figura 3 - Casamento de segmentos no espaço θ - s .

No modelamento efetuado na fase de treinamento, é determinado um vetor que vai do centro do segmento saliente do modelo ao centro de área do modelo.

Dado que o segmento saliente procurado foi encontrado na imagem (no espaço θ - s), efetua-se, no espaço cartesiano, uma rotação do ângulo de casamento no vetor associado ao segmento saliente encontrado e translada-se este vetor até que seu ponto inicial coincida com o centro do segmento de contorno da imagem. A localização do ponto final do vetor representa a localização do centro de área (posição) do objeto procurado na imagem, e, com a posição e inclinação do vetor, determina-se a orientação do objeto (figura 4).

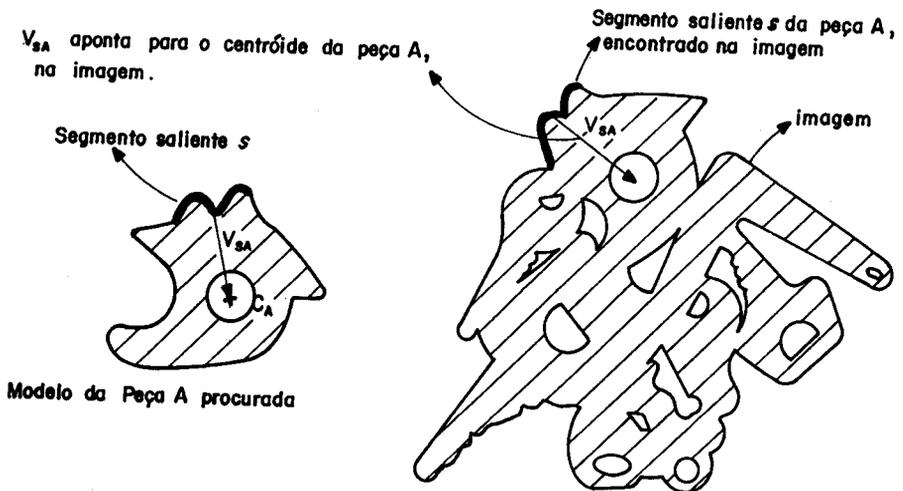


Figura 4 - Determinação da posição da peça procurada.

4. CONCLUSÃO

O sistema de visão artificial apresentado foi testado com sucesso em diferentes cenas contendo objetos isolados, objetos em contato e objetos em oclusão parcial.

Este sistema monta automaticamente, na fase de treinamento, a árvore de decisão binária que define qual a melhor ordem de teste dos atributos, dado um conjunto de peças apresentado ao sistema. Esta árvore de decisão, que pode ser montada para quaisquer conjuntos de peças, permite ganhos significativos de tempo, uma vez que define a maneira mais eficiente de busca.

Também os segmentos salientes são determinados automaticamente, dado um conjunto de peças apresentado na fase de treinamento.

Este sistema de visão está sendo utilizado na célula flexível de montagem da Divisão de Automação do LSI/DEE/USP.

BIBLIOGRAFIA

[1] RILLO, A.H.R.C. Sistema de visão binária, com reconhecimento de peças parcialmente oclusas. Dissertação de Mestrado apresentada na Faculdade de Engenharia Industrial de São Bernardo do Campo, 1989.

[2] PEREIRA, P.A.B. & VASSALLO, M. Módulo de aquisição de imagens coloridas. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO GRÁFICA E PROCESSAMENTO DE IMAGENS, 2, Águas de Lindóia, 1989. p.165-70.

[3] KITCHIN, P.W. & PUGH, A. Processing of Binary Images. In: PUGH, A. ed Robot Vision. Springer Verlag, 1983. p.21-42.

[4] HORN, B.P.K. Robot Vision. MIT Press, 1986.

[5] TURNEY, J.L. & MUDGE, T.N. & VOLZ, R.A. Recognizing partially hidden objects. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 2, Saint Louis, 1985. Proceedings. Washington, IEEE Computer Society Press, 1985. p.48-54.

[6] TURNEY, J.L. & MUDGE, T.N. & VOLZ, R.A. Recognizing partially occluded parts. IEEE Transactions on Pattern Analysis and Machine Intelligence, New York, 7(4): 410-421, July, 1985.