

USO DO FLUXO ÓTICO NA SEGMENTAÇÃO DE SEQUÊNCIAS DE IMAGENS

Jacques Facon

*Curso Pós-Graduação em Informática Industrial (CPGII)
Centro Federal de Educação Tecnológica do Paraná*

Av. 7 de Setembro 3165, 80230 Curitiba - Paraná

RESUMO : Apresenta-se um método de segmentação de uma cena de objetos em deslocamento ,cujos natureza e número são desconhecidos, a partir de uma sequência de imagens. A detecção do movimento será determinada a partir da noção do fluxo ótico e a interpretação final será formulada como um problema de classificação automática .

1. INTRODUÇÃO

O objetivo da segmentação de uma cena contendo objetos em deslocamento consiste em extrair zonas pertinentes e estáveis . No caso da segmentação de uma cena a partir de uma sequência de imagens, as principais dificuldades provêm da natureza da cena e do tipo de deslocamento de seus objetos. A necessidade de segmentar é ditada pela necessidade de distinguir e compreender os diferentes movimentos dos objetos da cena. A compreensão do movimento implica distinguir duas fases, a detecção do movimento e a avaliação dos parâmetros de deslocamento.

Esse artigo apresenta um algoritmo de segmentação de uma cena desconhecida por avaliação dos parâmetros de deslocamento dos objetos a partir de uma sequência de imagens , no caso onde o movimento escolhido dos objetos é uma translação 2D inteiramente contida no plano da cena . O número e o tamanho dos objetos são desconhecidos . A hipótese retida da natureza do movimento é a condição de rigidez do movimento [MITI-86] , que pode ser formulada como segue : " *um movimento é considerado como rígido se a distância entre cada par de pontos do objeto em movimento é invariável no tempo* ".A técnica proposta

leva ao fluxo ótico que será determinado por um processo de estabelecimento de correspondências entre duas imagens a partir de primitivas " pontos " . A parte 2 apresenta a seleção das primitivas usadas na determinação, na parte 3, do fluxo ótico . Na parte 4 é apresentado o método de segmentação da cena por análise do campo de fluxo ótico a partir da hipótese emitida acima .

2. ALGORITMO DE EXTRAÇÃO DE PRIMITIVAS

A extração de primitivas é baseada no algoritmo dos pontos de interesse de Moravec [MORA-77] . Um ponto da imagem é considerado como caracterizável para esse operador se e somente se ele é o centro de regiões de alta variância . Numa janela (5x5 nosso caso) centrada sobre esse pixel, são calculadas as variâncias para várias direções e o pixel central é substituído pela variância mínima . A imagem inicial é então resumida em uma imagem de pontos de interesse .

3. DETERMINAÇÃO DO FLUXO ÓTICO

Uma técnica consiste em estudar unicamente a estrutura e a natureza do movimento . A ferramenta apropriada é a noção de fluxo ótico [THOM-84] . O fluxo ótico representa o movimento aparente de objetos no campo visual humano de um observador [GIBS-57] . No domínio da visão por computador, ele representa o campo vetorial de deslocamento de entidades numa imagem.

Consideram-se duas imagens da seqüência de uma cena contendo um número desconhecido de objetos em movimento. O método de avaliação do fluxo ótico consiste em localizar e em seguir primitivas pequenas e identificáveis desses objetos nessa duas imagens. O operador de Moravec descrito acima é usado para segmentar as imagens . O resultado é então constituído de duas imagens novas de pontos. Esses dois conjuntos de pontos representam, em teoria, os pontos correspondentes ,nas duas imagens, dos objetos .

O campo de fluxo ótico é definido pelo estabelecimento

de uma correspondência entre os pontos de interesse dessas imagens. O objetivo do processo é fornecer pares de pontos correspondentes (P_i, Q_j) , P_i da primeira imagem e Q_j da segunda imagem. Cada $P_i Q_j$ representa o vetor deslocamento de um ponto. A hipótese de rigidez, expressa acima, permite afirmar que, no caso da translação, os pontos de um objeto criam vetores idênticos de deslocamento. A correspondência define então um conjunto de vetores que representa o deslocamento do objeto na sequência.

Para usar o conhecimento inicial contido nas imagens, a técnica aplicada é baseada no processo de relaxação de Barnard e Thompson [BARN-80] que leva em conta as disparidades entre as imagens. De uma forma iterativa, o método determina, em duas fases, a associação de pontos por avaliação de probabilidades. A primeira fase consiste em inicializar as probabilidades da correlação de dois pontos em função da qualidade de semelhança das vizinhanças dos pontos. Essa fase permite estabelecer um primeiro mapeamento refletindo as possibilidades reais de correspondência em função do critério de similaridade.

Essa primeira avaliação é melhorada na segunda fase a partir da propriedade de coerência do deslocamento. Essa propriedade, derivada da hipótese de rigidez, afirma, no caso de uma translação: "num ponto A , o deslocamento é coerente se os deslocamentos de pontos da vizinhança de A são idênticos". As probabilidades das associações são então atualizadas, de forma iterativa, em função da semelhança de deslocamento dos pontos vizinhos. A atuação para, quando as probabilidades são consideradas como suficientemente altas. O fluxo ótico relativo à evolução da cena é então constituído de vetores cujas probabilidades de correspondência são altas. Nosso critério de seleção é reter vetores cujas probabilidades são superiores a 0,7.

4. SEGMENTAÇÃO DO FLUXO ÓTICO

O campo vetorial obtido na etapa precedente dá uma representação geral de todos os deslocamentos dos objetos da

cena . Note-se que o número e o tamanho dos objetos na cena são desconhecidos . Segmentar o fluxo ótico consiste em decompor o conjunto de vetores em vários sub-conjuntos , cada sub-conjunto representando um deslocamento particular de um objeto . Considera-se o campo vetorial como um conjunto de dados a partir do qual deverão ser deduzidas classes . Cada sub-conjunto será interpretado como uma classe .

4.1 PARTICIONAMENTO INICIAL

Efetuamos a pesquisa de classes a partir de um método de classificação automática segundo o princípio das Nuvens Dinâmicas [DIDA-82] . Por isso, efetuamos uma mudança do sistema de eixos . Inicialmente no plano a imagem (O, X, Y) , o campo de vetores é representado no plano (O, Tx, Ty) onde Tx e Ty são os eixos de translação respectivamente em X e em Y .

Numa primeira etapa, os candidatos do conjunto inicial são agrupados , a partir de um critério de distância , aqui euclidiano, em volta de k núcleos aleatoriamente escolhidos . As k classes são representadas pelos centros de gravidades . Em volta desses são realizadas outras partições e o processo é reiterado até achar classes estáveis. Para evitar uma dependência das partições finais com os núcleos de partida, outras etapas semelhantes são realizadas a partir de outras escolhas aleatórias desses núcleos. A comparação das partições obtidas para cada escolha fornece classes homogêneas no plano (O, Tx, Ty) , cada classe representando um tipo de translação de um ou varios objetos da imagem . O pouco conhecimento da cena impede a fixação do número de classes . Então o algoritmo de particionamento foi desenvolvido de forma a detectar um número desconhecido de classes .

4.2 PARTICIONAMENTO FINAL

Na realidade, o processo descrito não é plenamente satisfatório porque a detecção de uma classe não implica sempre na detecção de um objeto da cena. A classe descoberta pode agrupar vários objetos. Esse caso pode ser interpretado de

duas formas :

- (1) - ele pode ser o resultado de uma interpretação errada e/ou de uma má avaliação . O processo de agrupamento pode ter reunido objetos cujos deslocamentos são quase iguais, considerando as translações como idênticas .
- (2) - objetos podem ter o mesmo deslocamento devido a uma relação de dependência geográfica, (estarem ligados, terem o mesmo suporte).

O interesse está na continuação do particionamento da classe . O particionamento inicial permite guardar no campo vetorial os vetores relativos à classe. Ele permite então selecionar os pontos de interesse de Moravec tendo participado na formação desses vetores. A ideia consiste, a partir da nuvem desses pontos retidos na primeira imagem, em explodir a classe por projeção sobre o eixo principal de inércia da nuvem .

Seja o caso de uma classe de translação agrupando em realidade dois objetos (figura I) . A projeção da nuvem de pontos correspondente sobre seu eixo principal de inércia fornece duas zonas de acumulação . A interpretação dessa projeção é realizada a partir de um histograma cumulativo . Esse histograma registra , de uma forma acumulativa, o número de projeções numa zona dada. Esse histograma fornece degraus onde alguns são maiores . Esses correspondem às zonas sem ponto para projetar . Esses degraus correspondem ao espaço livre entre os objetos na cena e constituem, no histograma , o critério de separação da classe em sub-classes .

Quando a cena contém vários objetos que se deslocam da mesma forma, pode acontecer que as projeções dos diferentes objetos sobreponham-se parcialmente . Nesse caso, o processo de análise pelo histograma é iterado para todas as sub-classes . Depois de uma primeira separação da classe original em duas sub-classes, cada uma delas é de novo analisada .O processo para , quando o critério de separação não é mais respeitado . A

determinação final dos parâmetros de translação é então melhorada por avaliação sobre cada classe final .

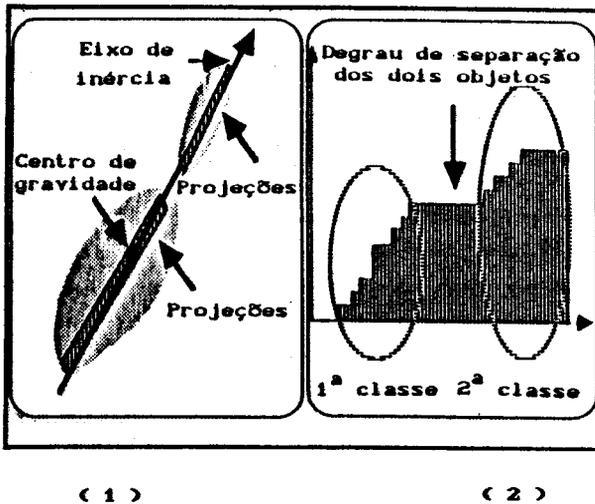


Fig 1 : Segmentação de uma classe contendo dois objetos

(1) projeção sobre o eixo de inércia

(2) ilustração do degrau de separação

Ilustramos o método em uma cena de laboratório que contém quatro objetos (figura II), dois deles tendo a mesma translação $T1 = (0 , 10)$, o terceiro deslocando-se de $T2 = (-10 , 0)$ e o último de $T3 = (10 , 0)$. A classificação automática forneceu três classes (figura II.4) e a análise por histograma permitiu dividir numa dessas em duas sub-classes (figura II.5). No total , a cena foi segmentada em quatro

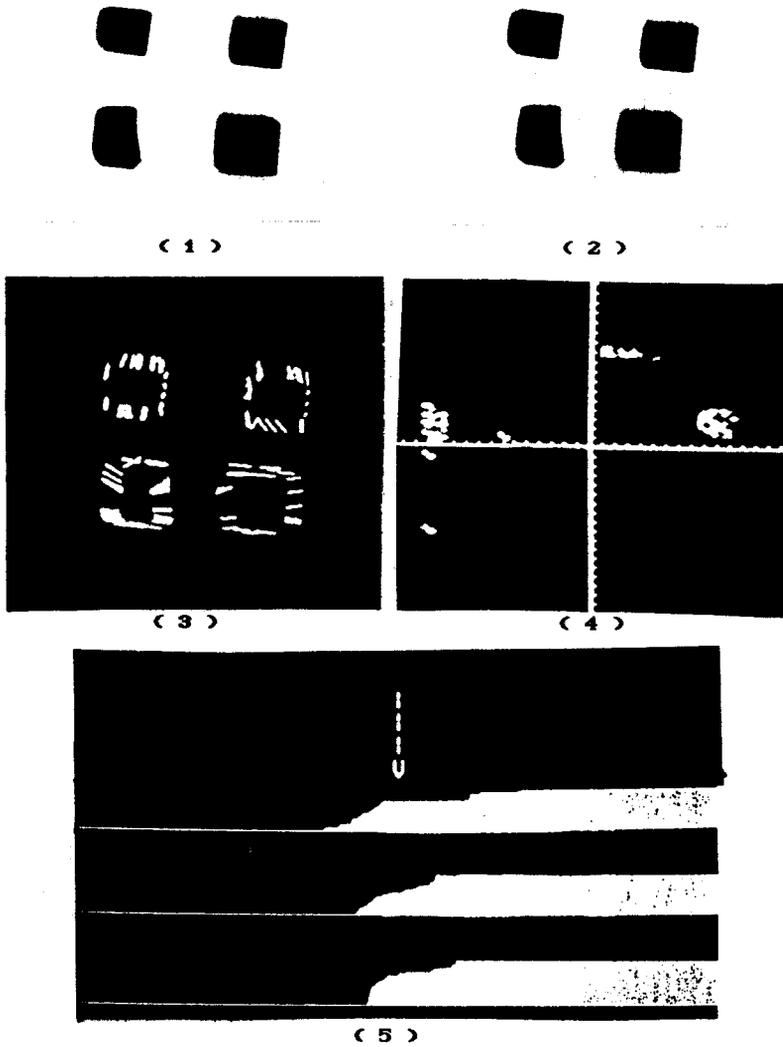


Fig II : Segmentação de uma cena

- (1) , (2) : duas imagens da seqüência
- (3) : fluxo ótico correspondente
- (4) : detecção das classes
- (5) : segmentação final

objetos , dois se trasladando (1 , 10) , um de (-12,1) e o último de (10 , 0) .

5. CONCLUSÃO

Mostramos que, apesar do pouco conhecimento disponível, a informação do movimento permitiu , no caso de uma translação, segmentar completamente e sem nenhuma ambigüidade uma cena totalmente desconhecida . Nenhuma restrição foi formulada sobre a natureza dos objetos e a amplitude do deslocamento . O algoritmo permitiu segmentar , de uma forma simples, cenas de objetos cujos deslocamentos eram muito diferentes e/ou parecidos ou iguais . Ele resolveu , entre outros, casos onde ambigüidades sobre a semelhança de deslocamento podiam aparecer.

A próxima etapa do processo será a segmentação de uma cena nos casos da rotação e , de uma forma geral, de um deslocamento qualquer .

6. REFERENCIAS

- [BARN-80] S.T. Barnard e W.B. Thompson , " Disparity analysis of images " , IEEE Transactions On Pattern Analysis And Machine Intelligence, vol PAMI 2, N°4 , Julho 1980, pp 333-340
- [DIDA-80] E. Diday , Elements d'analyse de données , Dunod, Bordas , Paris , 1982
- [GIBS-57] J.J Gibson, " Continuous perspective transformations and the perception of rigid motion, J Exp Psychol 54, 1957, pp 129-138
- [MITI-86] A. Mitiche , " On kineopsis and computation of structure and motion " , IEEE Transactions On Pattern Analysis And Machine Intelligence, vol PAMI 8, N°1 , Jan. 1986, pp 109-112
- [MORA-77] H.P Moravec , " Towards automatic visual obstacle avoidance " , Proc. 5th Int Joint Conf Artificial

Intell. Cambridge, MA, Agos. 1977 , p 584

[THOM -84] W.B. Thompson , K. M. Mutch e V.A. Berzins,
" Analyzing object motion based on optical flow ",
7th International Conference on Pattern
Recognition , Montreal, Canada, Julho 1984,
pp 791 - 794