

Registro espacial de imagens parcialmente sobrepostas: registros local e global.

JANDER MOREIRA

UFSCar-DC-Departamento de Computação
Via Washington Luiz, km 235
Caixa Postal 676
13560 São Carlos, SP, Brasil
ufscarc@brfapesp.bitnet

Abstract. This paper describes a method to register four images that cover a larger area. One image needs to have an overlapping region with an adjacent one. The process uses a two stage register. The first stage, called the local register stage, involves the matching of control points over two adjacent images and the register of these points. This provides a local matching point (the centroid of all control points) and the local register angle. This is done for each overlapping region. Since there are calculation errors in the local matchings, the four matching points may not agree with each other. Then the need of the second stage appears. This second step uses a minimization function to achieve the best position considering all of the four local matching points and their local angles. The function tries to force the points to agree and also tries to preserve the local register parameters.

1 Introdução

A montagem de mosaicos corresponde à determinação das posições relativas entre um dado conjunto de imagens. Há o pressuposto de que haja elementos nas diversas imagens que permitam a identificação de quais delas são adjacentes.

Este tipo de montagem é feito, normalmente de forma manual, no processo de fotomontagem em aerofotogrametria. Outras áreas de aplicação compreendem a montagem de fotos aéreas, com utilização em controle ambiental, poluição, preservação de recursos naturais etc.

O tratamento de montagem envolve etapas para sua realização, abrangendo basicamente dois estágios principais: o registro local e o registro global.

O registro local corresponde à realização de um *matching* para determinação dos parâmetros de registro entre duas imagens adjacentes. O conceito envolvido é a determinação de pontos de controle e sua localização em uma das outras imagens. Considerando-se que podem existir diferenças de translação e rotação entre as imagens, adota-se como medida de similaridade a correlação cruzada entre momentos invariantes [Hu (1962)]. De forma geral, as formas de similaridade usadas mais comumente são o coeficiente de correlação e a soma de diferenças absolutas [Svedlow (1978)]; entretanto nenhum destes métodos se mostra apropriado quando há diferenças de rotação.

O estágio de registro global considera o fato de

que o erro acumulado dos diversos posicionamentos locais pode gerar uma discrepância considerável em âmbito global. Apresenta-se, desta forma, a proposta de um esquema de ajuste global, baseado em um abordagem de Kamgar-Parsi [Kamgar-Parsi (1991)], utilizando o conceito de função de custo, como abordado por Fischler e Elschlager [Fischler (1973)].

2 Registro local

Goshtasby [Goshtasby (1986a)] define registro de imagens como "o processo de sobreposição de imagens da mesma cena". Outra definição sua é como "o processo de determinação das posições de pontos correspondentes em duas imagens da mesma cena" [Goshtasby (1986b)]. Embora a segunda definição já seja mais explícita que a primeira, o conceito de registro adotado neste texto é um pouco mais estenso. Entende-se, assim, como registro de imagens o processo de determinação da posição relativa entre duas imagens (e portanto entre seus pontos), sendo que ambas são diferentes enquadramentos de uma mesma cena mais ampla, considerando-se que elas podem ter em comum apenas uma pequena porção da cena original.

Desta forma, a determinação de qual imagem é adjacente a outra esbarra na limitação de que parte do conteúdo da cena deve estar presente em ambas. Em outras palavras, para que um dado par de imagens seja identificado como sendo de imagens adjacentes é necessário que haja uma área de sobreposição entre elas.

O processo de registro entre as duas imagens consiste em identificar pontos de controle presentes em ambas, permitindo assim a obtenção dos parâmetros que as diferenciam (em termos de deslocamentos ao longo dos eixos x e y e rotação relativos). Pontos de controle podem ser entendidos como pontos que identificam os mesmos objetos em cada imagem. Dados dois pontos de imagens distintas, ambos identificando um mesmo objeto presente em ambas, estes pontos são denominados pontos correspondentes.

Portanto, dadas uma imagem **A** e uma imagem **B**, ambas contendo apenas áreas parciais de uma mesma cena, sendo que existe uma região **R** comum a ambas, pretende-se estabelecer um conjunto de pontos de controle pertencentes a **R** e suas localizações nas duas imagens. O ponto de partida é a escolha arbitrária de pontos sobre a imagem **A**, para posterior determinação de seus correspondentes em **B**.

Definido um ponto sobre a imagem **A**, seleciona-se uma “janela” (ou subimagem) circular em torno dele. Esta janela vai ser, então, procurada, posição a posição, na imagem **B**. Esta busca faz a medida de similaridade entre a subimagem e a região da imagem **B** sob ela, para cada posição possível. A similaridade, como já mencionado, é obtida pela correlação dos momentos invariantes de cada subimagem. Dentre todas as posições, aquela que obtiver o valor mais elevado é considerada a “posição de *matching*”. Este processo é realizado para cada um dos pontos de controle.

Considerando-se que somente existem diferenças de translação e rotação entre o conjunto de imagens, dois pontos seriam suficientes para estabelecer os parâmetros (dx, dy, μ) , onde dx e dy são os deslocamentos em x e y , respectivamente e μ é o ângulo de rotação. Deve-se, entretanto, considerar que devido a diferenças na captura das imagens e inexatidão do próprio cálculo dos momentos, o *matching* tem certa imprecisão. Pode-se, na intenção de minimizar estes problemas, considerar um número maior (três ou mais) de pontos de controle, obtendo-se os parâmetros pela técnica de mínimos quadrados.

Foi considerado no processo de *matching* o uso de janelas circulares. Este fato justifica-se pelo seguinte ponto: não importa qual seja o ângulo de rotação, a área coberta por uma janela circular é sempre a mesma. Isto não ocorre com as tradicionais janelas retangulares, as quais quando rotacionadas passam a cobrir porções diferentes da cena. Estas considerações são bastante importantes para a obtenção de uma maior precisão nos cálculos de momentos.

Uma vez obtidos os parâmetros dx , dy e μ que diferenciam duas imagens, toma-se uma delas arbi-

trariamente como base, fazendo-se a rotação e translação necessárias da outra. Finalmente define-se o centróide dos pontos de controle como “ponto de registro” local, o qual será utilizado no processo de registro global.

3 Registro global

3.1 O problema angular

Considerando-se o grupo total de imagens, tem-se que aquelas adjacentes são conectadas através do ponto de registro. Entretanto, devido ao já citados problemas de precisão no *matching* e no registro local, discrepâncias entre os diversos pontos de registro podem ocorrer, de modo que não concordem entre si.

Adota-se, em analogia ao trabalho de Kamgar-Parsi [Kamgar-Parsi (1991)], o uso de uma função de custo para este ajuste global. Uma função similar já foi usada por Fischler e Elschlanger [Fischler (1978)] para o *matching* de figuras deformadas, usando uma analogia para função como se os elementos fossem interligados por “molas”, e sua distensão ou compressão causasse o aumento no “custo” do sistema.

Suponha-se quatro imagens que, após o estágio de registro local, apresentem uma disposição de 2×2 . Existem quatro pontos de registro local, identificados por I, J, K e L , nesta ordem. Fazendo-se a interligação destes pontos por retas, podem ser definidos os ângulos μ_{ij} (i, j pertencentes a $\{I, J, K, L\}$) como os que representam os ângulos entre cada par de imagens obtidos pelo registro local. Os ângulos β_i representam os ângulos formados entre as retas e o eixo horizontal (arbitrário), considerando-se o sentido anti-horário como positivo.

Os valores β_i são os que se deseja determinar. Como os valores μ_{ij} são calculados isoladamente e pode não haver concordância global, não necessariamente se tem, por exemplo, $\mu_{IK} = \beta_I + \beta_K$. O mesmo se aplica aos vértices restantes. De forma a se penalizar a função custo caso haja diferenças entre estes ângulos, introduz-se o termo

$$a_I(\beta_I - \beta_K - \mu_{IK})^2.$$

O fator a_I assume valor arbitrário, indicando o grau (ou peso) de importância dado ao termo. Uma vantagem sobre a forma do termo é que sua derivada é linear, facilitando a resolução do problema para os valores β_i .

Como cada vértice gera um termo semelhante ao descrito, a função de custo assume a forma

$$C = \sum_i (\beta_i - \beta_{i+1} - h_i \mu_{i,i+1})^2,$$

considerando-se que i assume ciclicamente os valores I, L, J e K , sendo que, se $i = I$, então $i + 1 = L$

e $i - 1 = K$, e assim por diante. O fator adicional h_i assume os valores -1, se $\beta_i < \beta_{i+1}$, ou 1, caso contrário.

A minimização da função custo é obtida derivando-a em relação a cada incógnita (os ângulos β_i) e igualando-se cada expressão a zero. Assim,

$$\frac{\partial C}{\partial \beta_i} = 2[(a_i - a_{i-1})\beta_i - a_i\beta_{i+1} - a_{i-1}\beta_{i-1} - (a_i h_i \mu_i, i + 1 + a_{i-1} h_i - 1 \mu_i - 1, i)].$$

Através da solução do sistema de equações lineares resultante obtém-se uma "otimização" para os ângulos.

Os parâmetros angulares a_i representam o grau em que um registro local $i, i + 1$ se deteriora quando a posição ideal (obtida no processo local) é perturbada em certa intensidade.

Supondo-se que o processo de registro local forneça um valor de similaridade $S_{i,i+1}$ para o ponto de registro $i, i + 1$, considera-se este valor como original. Se a posição for adulterada de um ângulo $d\mu$ (pequeno, como por exemplo 3° [Kamgar-Parsi (1991)]), obter-se-á uma nova medida de similaridade $S'_{i,i+1}$. Pode-se considerar a diferença $S_{i,i+1} - S'_{i,i+1}$ como o grau de degradação do registro local sob a alteração $d\mu$. Assim, um valor baixo para esta diferença indica baixa sensibilidade a mudanças, enquanto um valor elevado representa uma alta sensibilidade em relação à rotação.

Desta maneira, pode-se utilizar a_i proporcional a $S_{i,i+1} - S'_{i,i+1}$, ou simplesmente $a_i = S_{i,i+1} - S'_{i,i+1}$.

3.2 O problema translacional

Analogamente, deseja-se também manter a coerência de translação entre os pontos, já que cada ponto de registro existe em duas imagens e rotações destas imagens modificam suas posições. Desta forma uma função custo deve ser acrescida de termos que penalizem o distanciamento de pontos de controle correspondentes. Para cada ponto de registro deve-se ter, então, termos da forma

$$l_{xi}(x'_i - x''_i)$$

e

$$l_{yi}(y'_i - y''_i),$$

onde l_{xi} e l_{yi} são fatores que determinam o grau de penalização ao longo dos eixos x e y , respectivamente; (x'_i, y'_i) é o ponto de registro em uma imagem e (x''_i, y''_i) é o ponto de registro na outra.

A resolução, obviamente, é feita de maneira análoga à descrita anteriormente, ou seja, derivando-se em relação a cada incógnita e igualando-se as expressões resultantes a zero. Resolve-se, então, um sistema simples de equações lineares.

Os parâmetros l_{xi} e l_{yi} representam, de forma análoga ao raciocínio desenvolvido para os fatores angulares a_i , o grau de sensibilidade do registro local $i, i + 1$ sob uma diferença translacional dx ou dy , respectivamente.

Considerando-se que o ponto $i, i + 1$, com similaridade $S_{i,i+1}$, tenha sofrido alteração (pequena) dx , ao longo do eixo x , resultando em alteração da medida de similaridade para $S'_{i,i+1}$. Tomando-se $l_{xi} = S'_{i,i+1} - S_{i,i+1}$, assume-se que este representa a sensibilidade sob a translação dx . O mesmo se aplica ao longo do eixo y e a l_{yi} .

4 Considerações adicionais

A proposta para o processo de registro local baseia-se na utilização de momentos invariantes para o processo de *matching*. O cálculo de momentos é, porém, um processo demasiadamente custoso computacionalmente, envolvendo grande quantidade de cálculos. Considerando-se este aspecto, sugere-se a utilização da abordagem apresentada por Goshtasby [Goshtasby (1985)].

O principal conceito da abordagem é a adoção de dois estágios para a determinação da melhor posição de *matching*. Um primeiro estágio rápido (embora pouco preciso) seleciona candidatos para o segundo, que se utiliza de cálculos mais precisos e custosos para determinar o "melhor" dentre um número bem menor de posições.

Utiliza-se, dentro deste enfoque, o momento de ordem zero (que equivale à soma das intensidades dos pontos da janela) como seleção inicial, visto seu baixo custo computacional. As posições que obtiverem uma medida de similaridade "boa" (comparativamente a um limiar arbitrário) são selecionadas para o segundo estágio, onde momentos de maior ordem determinarão qual a melhor posição entre elas. Um ponto claro é que o primeiro estágio deve eliminar o maior número possível de candidatos, sem correr o risco de eliminar a posição correta. Um bom estudo sobre isto pode ser encontrado em um texto de Sitaraman e Rosenfeld [Sitaraman (1989)].

Considerações adicionais sobre *matching* em dois estágios podem ser encontradas também dois textos distintos de Rosenfeld e Vanderbrug [Rosenfeld (1977)] [Vanderbrug (1977)]. Uma abordagem hierárquica usando momentos invariantes de grande interesse consta de um texto de Wong e Hall [Wong (1978)].

O processo de registro global desenvolvido apresenta considerações para o caso do registro de quatro imagens. A extensão do processo para a montagem de mosaicos de maior magnitude está ainda sob análise, mas de forma geral considera-se que o esquema angular se aplica apropriadamente para cada conjunto de quatro imagens dispostas na configuração 2×2 . A parte que abrange o registro global considerando as tranlações dos pontos de registro correspondentes deve, entretanto, envolver todos os pontos simultaneamente. Evidentemente investigações mais profundas devem ser realizadas.

5 Conclusões

Uma metodologia para a montagem de imagens sobre uma região é apresentada. O conjunto de imagens adjacentes possuem uma região em comum, através da qual são feitos o *matching* e o registro local (para cada par de imagens), utilizando-se para isto a correlação de momentos invariantes.

Uma segunda fase do processo considera as imperfeições do registro local e faz o tratamento para um ajuste do registro em escopo global, considerando a redistribuição em termos de rotação e translação. Ambas, por meio de uma função de custo, tentam corrigir as imperfeições sem, entretanto, violar significativamente os registros locais.

O tratamento desenvolvido considera uma distribuição de quatro imagens distribuídas 2×2 . A extensão do processo para um conjunto de $m \times n$ imagens ainda se encontra sob análise.

A montagem de mosaicos é uma área de pesquisa bastante ampla e importante que não tem sido considerada com a ênfase apropriada. Trabalhos adicionais neste campo são necessários. Podem, assim, ser citados trabalhos relativos a fotomontagem e a tratamentos de qualidade nas intersecções das montagens, estes últimos já considerados por Milgram [Milgram (1975)] [Milgram (1977)].

6 Referências

- Hu, M.-K. Visual pattern recognition by moment invariants. *IRE Trans. on Inform. Theory*, **16**(2): 179-197, Febr 1962.
- Svedlow, M.; McGillem, C.D.; Anuta, P.E. Image registration: similarity measures and preprocessing method comparisons. *IEEE Trans. Aerosp. Electr. Syst.*, **14**(1): 141-149, Jan 1978.
- Kamgar-Parsi, B.; Jones, J.L.; Rosenfeld, A. Registrations of multiple overlapping range images: scenes without distinctive features. *IEEE Trans. Patt. Anal. Mach. Intell.*, **13**(9): 857-871, Sept 1991.

- Fischler, M.A. & Elschlanger, R.A. The representation and matching of pictorial structures. *IEEE Trans. Computers*, **22**(1): 67-92, Jan 1973.
- Goshtasby, A. Piecewise linear mapping functions for image registration. *Patt. Recogn.*, **19**(6): 459-466, 1986.
- Goshtasby, A.; Stockman, G.C.; Page, V.P. A region based approach to digital image registration with subpixel accuracy. *IEEE Trans. Geosc. Rem. Sens.*, **24**(3): 390-399, May 1986.
- Goshtasby, A. Template matching in rotated images. *IEEE Trans. Patt. Anal. Mach. Intell.*, **7**(3): 338-344, May 1985.
- Sitaraman, R. & Rosenfeld, A. Probabilistic analysis of two-stage matching. *Patt. Recogn.*, **22**(3): 331-343, 1989.
- Rosenfeld, A. & Vanderbrug, G.J. Coarse-fine template matching. *IEEE Trans. Syst. Man Cybern.*, pp. 104-107, Febr 1977.
- Vanderbrug, G.J. & Rosenfeld, A. Two-stage template matching. *IEEE Trans. Computers*, **26**(4): 384-393, Apr 1977.
- Wong, R.Y. & Hall, E.L. Scene matching with invariant moments. *Comput. Vis. Graph. Im. Proc.*, **8**: 16-24, 1978.
- Milgram, D.L. Computer methods for creating photomosaics. *IEEE Trans. Computers*, **24**(11):1113-1119, Nov 1975.
- Milgram, D.L. Adaptive techniques for photomosaicking. *IEEE Trans. Computers*, **26**(11): 1175-1180, Nov 1977.