

Segmentação de Vídeo Baseada em Movimento

Roberto Hugo Wanderley Pinheiro e Robson Cavalcanti Lins

Resumo—Este artigo apresenta uma modificação numa técnica de segmentação de vídeo, visando garantir melhores resultados e diminuir as dificuldades de segmentação em seqüências com movimentos lentos.

Palavras-Chave—Segmentação de Vídeo, Plano de Objeto de Vídeo, MPEG-4, Rastreamento de Objeto.

Abstract—This article presents an modification in a technique of video segmentation, aiming guarantee better resulted and diminish the segmentation difficulties in sequences with slow movements.

Keywords—Video Segmentation, Video Object Plane, MPEG-4, Object Tracking.

I. INTRODUÇÃO

O problema de segmentação do vídeo, tarefa fundamental para as aplicações no padrão MPEG-4 [1], consiste na extração dos objetos existentes em cada quadro da seqüência de vídeo. A abordagem mais comum para se iniciar a segmentação é baseada na diferença entre quadros, para encontrar os objetos em movimento. Um problema com essa metodologia, é que caso algum objeto deixe de se mover durante alguns quadros da seqüência, dificulta o seu rastreamento.

A técnica para segmentação de vídeo proposta por [2] mostrou-se eficiente quando a segmentação do primeiro quadro for precisa. Neste artigo, foi realizada uma modificação na referida técnica para suprir o problema gerado por uma segmentação falha do primeiro quadro.

II. ALGORITMO DE MEIER

Nesta seção descreve-se a técnica de segmentação de vídeo proposta por [2], para que após sua apresentação sejam mostrados os pontos modificados para a implementação realizada neste artigo.

Na técnica estudada, a imagem sofre uma série de operações para que o objeto possa ser extraído do quadro atual. Para que o algoritmo se desenvolva é necessário inicializar o modelo, seguir atualizando durante a seqüência do vídeo para então extraí-lo e encontrar todos os quadros segmentados. Nas subseções seguintes serão apresentados estes passos do algoritmo.

A. Inicialização do modelo

Para dar início à execução do algoritmo é necessário gerar um modelo inicial para que ele seja utilizado como parâmetro da segmentação do primeiro quadro e gerar um modelo do

segundo quadro, para segmentar o segundo e assim consecutivamente durante toda a seqüência de vídeo.

O primeiro modelo é gerado pela diferença entre os dois primeiros quadros da seqüência. O resultado será uma imagem de difícil visualização, mas ao realizar uma limiarização transformando-a em uma imagem de apenas duas cores (preto e branco) será possível notar as partes que se movem durante esses dois quadros, considerado o objeto ou parte dele. Sobre esta imagem aplica-se um algoritmo de afinamento de bordas e o modelo inicial está pronto para ser utilizado.

B. Atualização do modelo

Na atualização do modelo, serão recebidos como parâmetros para sua execução o modelo inicial e a imagem da borda.

A imagem da borda, gerada pelo Operador Canny, será o principal instrumento da segmentação, já que ela delimitará as bordas do objeto durante o processo.

Alguns passos são necessários para atualização do modelo, eles serão listado à seguir:

1) *Componentes Rápidos e Lentos*: Para tentar evitar que os objetos que se movem pouco se percam durante a seqüência de vídeo, é realizada uma diferente abordagem para estes.

A imagem da borda é dividida em dois componentes, um representa os que se movem rapidamente e outro os que se movem lentamente, que serão concatenados ao final das duas operações. Para tal operação utiliza-se a Distância de Hausdorff e Rotulação de Componentes Conectados.

2) *Distância de Hausdorff*: Proposta por Huttenlocher [3], a distância de Hausdorff, tem como objetivo comparar imagens binárias. Essa comparação será feita entre o modelo inicial e a imagem da borda. Com isso, serão detectados os pixels da imagem da borda que se encontram a uma distância dos pixels do modelo inicial.

A implementação da distância de Hausdorff foi realizada via a Transformada da Distância (DT) que define pequenas máscaras contendo aproximações inteiras de distâncias de uma pequena vizinhança. Existem duas máscaras que podem ser utilizadas: Chamfer 3-4 e Chamfer 5-7-11. A técnica em estudo utilizou a máscara Chamfer 5-7-11 por gerar melhores resultados. Ao aplicar a DT será obtido os componentes rápidos do quadro atual.

3) *Rotulação de Componentes Conectados*: A rotulação é necessária para que possa encontrar os prováveis objetos do quadro, a técnica é aplicada sobre um quadro de bordas e limiarizado, sendo então uma rotulação dos pixels da borda que estão conectados entre si.

Nesta fase do algoritmo, será encontrado os componentes lentos que serão obtidos encontrando as conexões entre o modelo inicial e a imagem da borda.

Unindo-se os dois componentes já encontrados será obtido a imagem da borda que representa o objeto.

4) *Remoção de Fundo Ruidoso*: Para evitar ruídos impostos por fundos confusos ou complexos, é aplicada uma técnica para remover os ruídos que se apresentam na imagem da borda.

Basicamente é realizado um histórico de toda a seqüência de vídeo, até o momento para que seja montada uma contabilização sobre cada pixel. Aquele que ultrapassar um determinado número de aparições, será considerado fundo e removido.

C. Extração dos Objetos

O processo de extração do objeto é uma aplicação simplista de varredura em toda a imagem nos eixos x e y.

Nesse estágio a imagem estará segmentada possuindo apenas pixels que pertençam a borda do objeto. O algoritmo irá passar por todas as linhas da imagem e do primeiro ao último ponto da borda de cada linha ele colocará os respectivos pixels da imagem original. O mesmo se aplicará para as colunas, formando no final o objeto segmentado.

III. MODIFICAÇÕES REALIZADAS

A modificação realizada consiste na junção dos componentes rápidos e lentos, para uma vez unidos serem detectados em conjunto ao invés de concatenados posteriormente, para isso foi realizada uma modificação na técnica de Rotulação de Componentes Conectados.

A técnica de rotulação fará uso da imagem da borda e a imagem gerada pela DT. Todos os pixels da imagem gerada pela DT serão considerados como pertencentes à mesma região, mesmo não estando conectados, essa implementação se deve ao fato de poder encontrar todas as bordas que se conectam por meio da parte que se move da imagem, sendo possível detectar os movimentos suaves.

IV. RESULTADOS

As simulações foram feitas com o uso dos componentes de luminância para imagens das seqüências clássicas: *Claire* e *Tênis*. A seqüência *Tênis* possui movimentos bruscos, enquanto a *Claire* possui movimentos suaves, portanto foram analisados tipos diferenciados de movimento. O tamanho utilizado em cada imagem foi 240 x 240 pixels.

Os resultados mostram a segmentação, de apenas um quadro, realizada sobre as seqüências descritas. Na figura 1 apesar de ser uma seqüência composta por movimentos suaves e lentos, foi possível extrair com perfeição o objeto presente na cena. Devido ao fundo pouco ruidoso e constante não houveram problemas na segmentação. Já na figura 2, mesmo possuindo movimentos bruscos e notáveis a segmentação gera ruídos. Esses ruídos provêm justamente da grande movimentação entre dois quadros, deixando resquícios de movimentos, ou ruídos da própria imagem.

V. CONCLUSÕES

A modificação na técnica surtiu um efeito positivo sobre a seqüência *Claire*, que possui movimentos lentos, mas gerou ruídos na seqüência *Tênis*, que é uma seqüência de movimentos rápidos.



Fig. 1. Segmentação do 2º quadro da seqüência Claire.

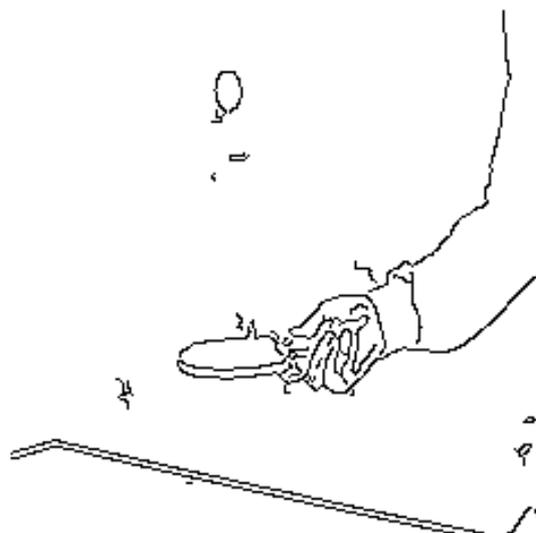


Fig. 2. Imagem da borda do 3º quadro da seqüência Tênis.

Conclui-se que a técnica possui bons resultados para seqüências lentas, que são comumente de difícil segmentação devido a detecção do movimento ser complexa. Porém, para as seqüências mais rápidas é necessário o acréscimo de algum componente que comprometa menos a segmentação.

AGRADECIMENTOS

Os autores agradecem à Universidade Católica de Pernambuco pelo suporte a pesquisa.

REFERÊNCIAS

- [1] T. Sikora, "The MPEG-4 Video Standard Verification Model", *IEEE Transaction Circuits and Systems for Video Technology*, v.7, n.1, p.19-31, 1997.
- [2] T. Meier e K. N. Ngan, "Automatic Segmentation of Moving Objects for Video Object Plane Generation", *IEEE Transactions on Circuits and System for Video Technology*, v.8, n.5, 1998.
- [3] D. P. Huttenlocher, G. A. Klanderman e W. J. Rucklidge, "Comparing images using the Hausdorff Distance", *IEEE Transaction Pattern Analysis and Machine Intelligence*, v.5, n.9, 1993.