

Filtragem Morfológica para Redução do Ruído na Detecção de Movimento em Sistema Automático de Vigilância

Simara Sonaglio
CST em Sistemas de Telecomunicações, IF-SC
São José, Santa Catarina, 88103-310, Brasil

e

Marcos Moecke
Departamento de Telecomunicações, IF-SC
São José, Santa Catarina, 88103-310, Brasil

RESUMO

Este trabalho apresenta um sistema automático de detecção de movimento baseado apenas em sinais de vídeo. A informação de vídeo é capturada de uma câmera de baixo custo do tipo webcam. Neste sistema a detecção de movimento é feita pela subtração de fundo e a redução de ruídos é feita através da filtragem morfológica. Para facilitar a análise por humanos, e economizar espaço de armazenamento são apenas armazenados os quadros com movimento. O instante no qual ocorreu cada evento é marcado através da sobreposição de um texto em cada quadro. A identificação do objeto causador do movimento é realizada pela segmentação desse objeto usando-se uma moldura para demarcá-lo no vídeo armazenado. Os resultados obtidos mostram que o sistema é eficiente na detecção de movimentos, evitando alarmes falsos devido à variação da iluminação, sombras, alteração de fundo e movimentos causados pelo vento.

Palavras chave: Detecção de movimento, vídeo, monitoramento de ambientes, filtragem morfológica, ruído.

1. INTRODUÇÃO

Recentemente está ocorrendo uma grande expansão no mercado de equipamentos com a capacidade de detectar o movimento. Esses equipamentos podem detectar movimento através de sensores externos ou pela própria imagem. Nestes sistemas a observação e interpretação humana foram substituídas por algoritmos computacionais automatizados e inteligentes [1], [2].

A detecção de movimento através de sinais de vídeo e a segmentação da imagem têm forte apelo de uso, seja em sistemas de vigilância eletrônica ou em outras aplicações como controle de acesso, visão robótica e sistemas de auxílio à direção de veículos [1]. A

maioria desses sistemas segmenta a imagem, separando o objeto em movimento do fundo estático com o objetivo de detectar o movimento na cena. O processo de detecção do movimento normalmente envolve tanto a modelagem do ambiente, que consiste na representação da imagem de fundo, como a segmentação do objeto em movimento e a sua classificação [1]. Durante o processo de detecção de movimento podem ocorrer falsos alarmes de detecção fazendo com que movimentos que não aconteceram sejam detectados e que movimentos existentes na cena passem não detectados. A produção destes falsos alarmes depende das condições de aquisição da imagem, resolução da câmera, taxa de quadros, ruído de aquisição, e das condições do ambiente monitorado como iluminação, sombras, vento, reflexos [3]. O acompanhamento da trajetória do objeto segmentado e a gravação das imagens para posterior análise são opções que também são freqüentemente implementadas. Um sistema baseado apenas em sinais de vídeo com intuito de reconhecer padrões de movimentação de um ser humano é apresentado em [4], no qual a detecção de movimento é realizada através da técnica de subtração de fundo e da filtragem morfológica. Este sistema utiliza uma câmera de com sensor CCD de maior custo que as câmeras empregadas no presente trabalho. Além disso, o sistema apresentado em [4] tem uma maior complexidade devido ao reconhecimento de padrões feito após a detecção de movimento.

Este artigo visa descrever o desenvolvimento de um sistema de vigilância denominado vigilância por detecção automática de movimento (VDAM). O VDAM utiliza o método denominado detecção de movimento por filtragem morfológica (DMFMorf). O sistema VDAM permite a gravação em mídia de armazenamento digital das imagens relevantes das áreas supervisionadas, eliminando a necessidade de acompanhamento permanente de um observador humano. Para facilitar a análise posterior de eventos e

a identificação dos objetos é acrescentada a informação de data e hora do evento, e uma moldura em torno dos objetos causadores do movimento.

Na Seção 2 é apresentado o sistema VDAM, sendo abordada a aquisição das imagens, a detecção do movimento, a atualização de fundo, o armazenamento em arquivos, e controle da exposição à luz da câmera. Na Seção 3 são apresentados os equipamentos utilizados para o desenvolvimento deste sistema, bem como os resultados obtidos. Finalizando o artigo são apresentadas as conclusões mais relevantes e algumas propostas de trabalhos futuros.

2. SISTEMA PROPOSTO

O sistema de vigilância por detecção automática de movimento (VDAM) tem como objetivo realizar a detecção de movimento baseado apenas em sinais de vídeo, sendo capaz de eliminar informações irrelevantes captadas na aquisição da imagem, como ruído de aquisição e condições do ambiente como iluminação, sombras, vento, reflexos.

O diagrama de blocos da Figura 1 mostra as etapas deste sistema. O VDAM detecta movimento através do método detecção de movimento por filtragem morfológica (DMFMorf) e a alteração de fundo é realizada pelo algoritmo de detecção de fundo por blocos (DFB). No VDAM é inserida uma moldura em torno do objeto causador do movimento para facilitar a sua identificação, sendo utilizado para isso o algoritmo inserção de molduras múltiplas (IMM). As imagens são armazenadas com o padrão de compressão de vídeo MPEG4 através do codec XVID.

A. Aquisição da imagem

O sinal de vídeo consiste basicamente em uma seqüência de imagens reproduzidas continuamente a uma taxa de quadros por segundo (fps). Quanto maior a taxa melhor será a percepção de movimento pelo sistema visual humano. Cada quadro de imagens é representado por uma matriz de *pixels*, que pode ser bidimensional no caso de imagens preto e branco ou tridimensional para imagens coloridas [5].

No sistema VDAM, a informação de vídeo é capturada com uma câmera de baixa resolução do tipo webcam.

A câmera tem sensores CMOS de baixo custo, e por isso apresenta um forte ruído de aquisição, principalmente em ambientes pouco iluminados. Para reduzir este ruído e diminuir a detecção de falsos movimentos, as imagens adquiridas passam por uma filtragem mediana antes de serem processadas. Esta filtragem é aplicada nas componentes de cor Y, Cb e Cr tanto da imagem de fundo quanto da imagem atual.

B. Detecção de movimento

A detecção de movimento através de sinais de vídeo tem como objetivo básico detectar o movimento em ambientes com predominância de cenas estáticas, ou seja, em ambientes em que normalmente não ocorre movimentação [6], [7]. A implementação da detecção de movimento pode ser feita através de duas técnicas principais, a subtração de fundo e a diferença temporal entre imagens [1]. A subtração do fundo consiste na diferença espacial entre o modelo do ambiente e a imagem atual captada. A diferença temporal entre imagens consiste na subtração dos *pixels* de quadros sucessivos.

O método de detecção de movimento DMFMorf proposto neste trabalho utiliza a técnica de subtração de fundo. Esse método é ilustrado na Figura 2, sendo constituído das seguintes etapas:

Etapa 1: Redução do ruído de aquisição da câmera através da filtragem mediana da imagem atual e da imagem de fundo, em blocos de 3×3 *pixel*. A filtragem é realizada por:

$$y(k,l) = \frac{1}{9} \sum_{j=k-1}^{k+1} \sum_{i=l-1}^{l+1} x(i,j) \quad (1)$$

onde $x(i,j)$ são os *pixels* originais e $y(k,l)$ são os *pixels* filtrados.

Etapa 2: Subtração *pixel-a-pixel* da imagem de fundo da atual, resultando na matriz de diferenças absolutas de *pixels*.

$$\Delta(i,j) = |y_{\text{Fundo}}(i,j) - y_{\text{Atual}}(i,j)| \quad (2)$$

$\forall i, j \in \text{Imagem}$

onde $y_{\text{Fundo}}(i,j)$ são os *pixels* da imagem de fundo, $y_{\text{Atual}}(i,j)$ são os *pixels* imagem atual.

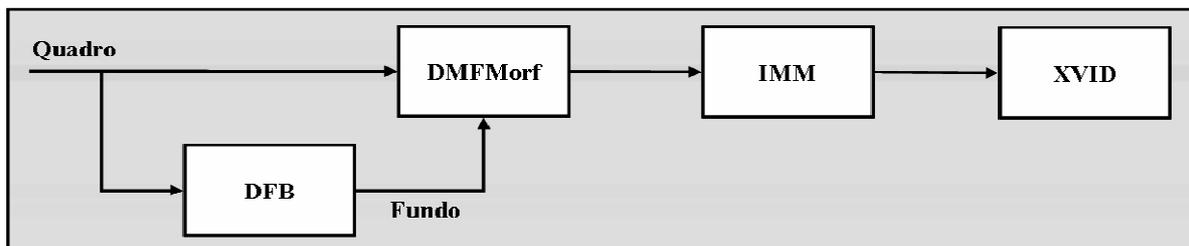


Figura 1 – Diagrama de blocos do sistema VDAM.

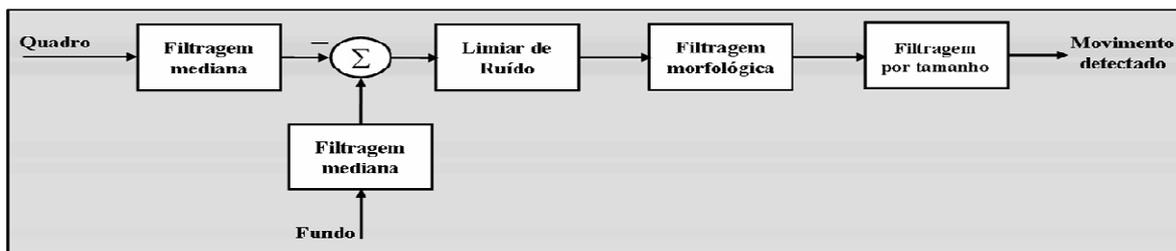


Figura 2 - Diagrama de blocos do DMFMorf.

Etapa 3: Obtenção da imagem binarizada, a partir da comparação dos *pixels* da matriz de diferenças com os limiares de ruído determinados.

$$B(i, j) = (\Delta(i, j) > \text{Limiar}) \quad (3)$$

$$\forall i, j \in \text{Imagem}$$

Etapa 4: Realização da filtragem morfológica através das operações *Abrir* e *Fechar* para eliminar ruídos na imagem binarizada B e preencher o contorno dos objetos que eventualmente não ficaram bem segmentados. Este preenchimento visa juntar partes próximas de um objeto que foram separadas durante a segmentação. A operação *Abrir* realiza uma erosão seguida de uma dilatação e é utilizada na eliminação de ruídos. A operação *Fechar* faz uma dilatação seguida de uma erosão e é utilizada para preenchimento [5].

A Figura 3 apresenta um exemplo de uso dessas operações. Na imagem (a) é apresentada a imagem binarizada com ruído. Em (b) após a operação *Abrir* o ruído existente em (a) foi eliminado. Na imagem (c) é apresentado o resultado para a operação *Fechar*, onde o contorno do objeto da imagem foi melhorado.

Etapa 5: Na filtragem por tamanho os segmentos com área de *pixel* menor que 100 *pixels* são excluídos.

O algoritmo DMFMorf para a detecção de movimento é apresentado na Figura 4.

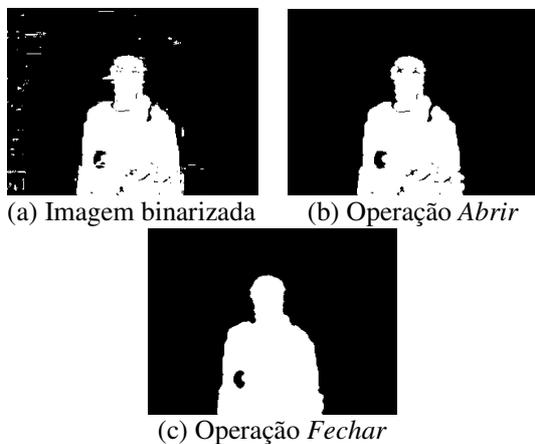


Figura 3 - Exemplo de utilização da filtragem morfológica.

Laço infinito

- **Aquisição das imagens**
- **Conversão de espaço de cor:** $YCbCr = rgb2ycbcr(\text{Atual})$
- **Filtragem mediana:** $y(k, l) = \frac{1}{9} \sum_{j=k-1}^{k+1} \sum_{i=l-1}^{l+1} x(i, j)$
- **Obtenção da imagem de fundo:** $y_{\text{Fundo}} = DFB(y_{\text{Atual}})$
- **Subtração de fundo:** $\Delta(i, j) = |y_{\text{Fundo}}(i, j) - y_{\text{Atual}}(i, j)|$
 $\forall i, j \in \text{imagem}$
- **Aplicação dos limiares de ruído:** $B(i, j) = \Delta(i, j) > \text{Limiar}$
 $\forall i, j \in \text{imagem}$
- **Filtragem morfológica:** $\text{ObjSeg} = \text{fechar}(\text{abrir}(B))$
- **Filtragem por tamanho:** $\text{ObjSegFinal} = \text{ObjSeg} > 100$

Fim laço infinito

Figura 4 – Algoritmo DMFMorf.

C. Detecção de alteração no fundo

O fundo é uma imagem estática utilizada como referência na detecção de movimento. A adequada detecção de movimento depende de uma imagem de fundo correta, por isso é necessário modelar o ambiente de três dimensões em uma imagem de duas dimensões.

Em um sistema de vigilância a imagem de fundo necessita ser constantemente atualizada. Essa atualização deve compensar as variações de iluminação e alterações da cena produzidas pela mudança de objetos no leiaute do ambiente, que uma vez trocados de posição, e tornados estáticos, não devem mais ser considerados como movimento pelo algoritmo de detecção [8].

A Figura 5 ilustra um exemplo de alteração de ambiente no qual um objeto que estava no ambiente (Cenário1) foi retirado. Quando a mudança de ambiente persistir por algum tempo, é necessário tratá-la adequadamente adotando um novo modelo de ambiente para representar o novo cenário (Cenário2), evitando a gravação de sucessivos quadros em uma cena estática.



Figura 5 - Exemplo de alteração de ambiente.

No VDAM, a detecção de alteração de fundo é feita através da técnica proposta DFB que usa a diferença temporal entre a imagem atual e anterior (ver Figura 6). A divisão da imagem de fundo em 16 blocos visa permitir que o fundo seja atualizado em cenas nas quais haja movimento em parte da imagem. Quando é detectada a ausência de movimentos em um dos blocos o fundo correspondente ao mesmo é atualizado, evitando assim detectar erroneamente objetos que foram removidos como sendo um objeto em movimento.

D. Gravação da imagem

O sistema VDAM visa principalmente à eliminação da análise contínua de cenas por observadores humanos, mas não dispensa a análise posterior dos eventos sempre que se fizer necessário. Para tal, as imagens nas quais são detectados movimentos significativos são armazenadas em arquivo. A informação de data e hora é sobreposta aos quadros armazenados para permitir localizar no tempo o evento. Além disso, os objetos causadores de movimento são emoldurados, para facilitar a sua visualização.

O sistema proposto além de armazenar todos os quadros com movimento, também armazena as alterações de imagem de fundo e um quadro anterior e posterior ao movimento. Dessa forma é possível verificar as alterações que ocorreram na cena.

Para facilitar a identificação dos arquivos, o seu nome é formado pela data e hora em que foi armazenado o primeiro quadro, e através da organização das pastas, os arquivos de vídeo são separados por mês e ano. Cada arquivo de vídeo tem no máximo 900 quadros.

O sistema atualmente necessita em média 2,6 kbytes para armazenar cada quadro com resolução de

320x240 *pixels*. Nas situações em que ocorre movimento contínuo na cena as imagens são armazenadas a uma taxa de 3 quadros por segundo.

E. Ajuste da exposição à luz da câmera

Em uma câmera é fundamental controlar a quantidade de luz que passa pela abertura de seu diafragma e atinge o sensor de imagens. Neste trabalho foi utilizada uma câmera do tipo webcam, prevista para uso preferencialmente em ambientes internos, nos quais a iluminação é controlada. Este tipo de câmera apresenta problemas quando a exposição à luz da mesma é muito baixa (ambientes escuros) e quando a exposição é muito alta (ambientes externos muito claros).

Para controlar a exposição à luz, foi desenvolvido um algoritmo denominado Variação Contínua da Exposição à Luz (VCEL). Este algoritmo calcula o valor médio dos *pixels* capturados, e ajusta o tempo de exposição de modo a obter valores médios na faixa de 100 a 200.

Se a média for abaixo de 100, a imagem é considerada muito escura e a exposição à luz é aumentada, por outro lado, se a média está acima de 200, a exposição à luz é reduzida, pois a imagem está muito clara. Esse procedimento garante a qualidade da imagem a ser processada, independentemente das variações de iluminação do ambiente que ocorrem durante a captura.

O algoritmo VCEL para o controle da exposição à luz da câmera é mostrado na Figura 7.

3. EQUIPAMENTOS E RESULTADOS

Para o desenvolvimento do sistema são utilizados: i) câmera de vídeo do tipo webcam A4TECH modelo PK-333MB, com resolução entre 160x120 e 1280x960 e taxa de quadros por segundo de 6 a 30fps; ii) computador com processador Intel Pentium Dual Core de 1,6GHz, com memória de 2GB, sistema operacional Windows XP e monitor de LCD de 17" com resolução de 1280x1024 *pixels*; iii) software MATLAB para o desenvolvimento dos algoritmos de detecção de movimento, segmentação da imagem e gravação do vídeo.

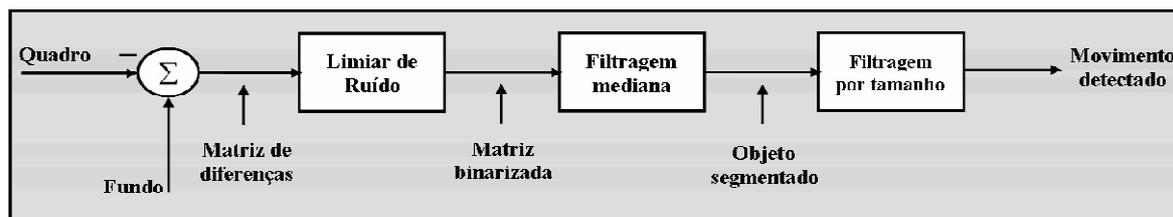


Figura 6 - Diagrama de blocos da detecção de alteração de fundo do algoritmo DFB.

Laço infinito

- Aquisição das imagens
- Cálculo do valor médio dos pixels:

$$\bar{x} = \frac{1}{320 \times 240 \times 3} \sum_{i=1}^{320} \sum_{j=1}^{240} \sum_{k=1}^3 x[i, j, k]$$

Se $\bar{x} \leq 100$

- Aumenta a exposição à luz

Senão se $\bar{x} \geq 200$

- Diminui a exposição à luz

Fim se \bar{x}

Fim Laço infinito

Figura 7 - Algoritmo VCEL.

O sistema VDAM mostrou bons resultados quanto à segmentação do objeto e redução de ruídos. Como pode ser visto na Figura 8, em uma sequência de 50 quadros, onde havia em cada quadro apenas um objeto em movimento, o sistema não detectou corretamente o movimento em apenas 4 deles. O número máximo de objetos encontrados em uma imagem foi de dois. Este erro ocorreu pela não eliminação de ruído ou pela má segmentação do objeto.

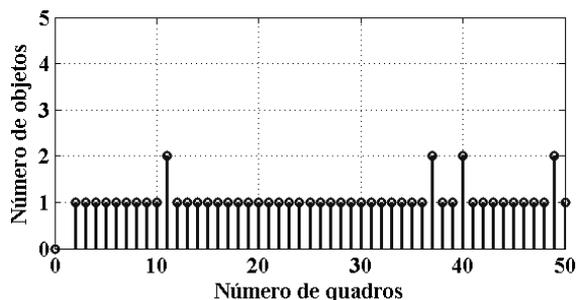


Figura 8 – Ilustração da detecção do número de objetos em uma sequência de vídeo.

O sistema foi testado em ambiente interno e externo. Em ambientes internos foram obtidos bons resultados quanto à diminuição de alarmes falsos. Neste caso, devido à iluminação ser constante e normalmente proveniente do teto, a variação da luminosidade e a projeção de sombras é pequena. Apesar disso, em algumas situações podem ocorrer falsas detecções de movimento, por exemplo, acendimento ou desligamento de uma lâmpada, e a abertura ou fechamento de alguma porta ou janela. Também podem ocorrer modificações no ambiente devido à retirada ou colocação de objetos no ambiente, que são detectadas pela técnica DFB e incorporadas ao modelo do ambiente.

Em ambientes externos os resultados obtidos foram satisfatórios, porém com maior número de alarmes falsos. O desempenho do sistema foi reduzido devido ao aumento do número de objetos falsos detectados na cena. Ao mesmo tempo a troca da imagem de fundo tornou-se mais freqüente devido às variações de luminosidade durante o dia ou presença de nuvens

obscurecendo a luz solar. Por outro lado, no cenário externo podem ocorrer movimentos que o sistema não deseja detectar, tais como a movimentação de folhas pela ação do vento. Outro problema é a ocorrência de reflexos de luz em superfícies lisas que podem ser detectadas como movimentos falsos. Finalmente, também podem ocorrer alterações temporárias do cenário devido à mudança de objetos visuais tais como: móveis, abertura de janelas e portas, veículos, etc., que devem ser incorporadas ao modelo de fundo.

4. CONCLUSÕES E TRABALHOS FUTUROS

Neste artigo foi proposto um sistema de vigilância por detecção automática de movimento - VDAM. Foram usadas imagens adquiridas através de câmeras de baixo custo, com o objetivo de viabilizar o desenvolvimento de produto de baixo custo que possa ser utilizado em ambientes públicos, principalmente escolas municipais e estaduais. Nestes ambientes a existência de vigilância tende a reduzir os pequenos delitos e a depredação do patrimônio.

A visualização e análise das imagens por operadores humanos são facilitadas pelo sistema, através do armazenamento das imagens com movimentos significativos. Essa abordagem resulta em um tamanho de arquivo menor, e ao mesmo tempo reduz a fadiga dos vigilantes encarregados de observar diversos monitores de imagem. Usando esse sistema, não é necessário acompanhar constantemente as imagens, mas apenas analisar os eventos de movimento detectados e sinalizados pelo sistema.

A ocorrência de falsos alarmes devido à movimentação de vegetação, sombras, variação de luminosidade ambiente foi significativamente reduzida pelos algoritmos de detecção e troca de fundo propostos. Nos testes realizados a maioria dos movimentos foi detectada com sucesso tanto em ambientes internos como externos. As falhas de detecção ocorreram quando as pessoas estavam muito distantes da câmera, ou utilizavam roupas que se assemelhavam ao fundo da imagem. O uso de câmeras de maior resolução provavelmente reduziria as falsas detecções. Não obstante, o sistema com baixa resolução desenvolvido é útil no auxílio a vigilância de ambientes, pois as imagens armazenadas são posteriormente analisadas por observadores humanos.

No sistema desenvolvido as alterações de iluminação principalmente de ambientes externos são compensadas pelo algoritmo VCEL. O VCEL detecta quando a imagem está muito escura ou

clara e ajusta o tempo de exposição à luz melhorando o contraste da imagem e minimizando os efeitos das variações de iluminação do ambiente.

O protótipo do sistema VDAM foi desenvolvido em Matlab devendo ainda ser implementado em linguagem de programação compilada e/ou DSP para a obtenção de um produto final com maior eficiência quanto à velocidade de processamento. Como continuidade deste trabalho, propõe-se que seja realizado:

- Implementação do sistema VDAM em uma plataforma de hardware do tipo processador de sinais digitais (DSP) ou arranjo de portas programáveis em campo (FPGA).
- Desenvolvimento de uma interface gráfica que permita recuperar e comentar as imagens gravadas, bem como acompanhar visualmente o sistema.
- Validação completa do sistema com a realização de testes de campo, nos quais diversos movimentos de pessoas precisam ser analisados por observadores humanos e pelo sistema proposto.
- Uso das imagens detectadas com movimento para sua transmissão através da rede local e internet [9] e através da rede celular pelo sistema GSM/EDGE [10].
- Controlar o movimento da câmera a partir das informações obtidas da imagem, para acompanhar os objetos em movimento de modo automático [11].
- Desenvolvimento de um algoritmo que permita a detecção de sombras.
- Desenvolvimento de um algoritmo que permita o registro de passagem e identificação de pessoas através da detecção do rosto.

5. REFERÊNCIAS

- [1] L. Wang; W. Hu; T. Tan; S. Maybank, “A survey on visual surveillance of object motion and behaviors”, IEEE Trans. on Systems, Man and Cybernetics, Vol. 3, 2004, pp. 334-352.
- [2] C. Cedras; M. Shah, “Motion-based recognition: a survey”, Image Vision Comput. Vol. 13, 1995, pp.129–155.
- [3] D. Toth; T. Aach; V. Metzler, “Illumination – invariant change detection”, 4th IEEE Southwest Symposium Image Analysis and Interpretation, Proceedings, 2000, pp. 3-7.
- [4] A. Basu; M.K. Mandal; M. Singh, “Human Activity Recognition Based on Silhouette Directionality”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 18, No. 9, 2008.
- [5] W.K. Pratt, Digital image processing, New York: Pks Inside, 3ª ed., 2001, pp. 735.

- [6] B.J. Jones, “Low-cost outdoor video motion and non-motion detection”, 29th Annual 1995 International Carnahan Conference on Security Technology, IEEE, 1995, pp. 376-381.
- [7] R.H. Pinheiro; R.C. Lins, “Segmentação de vídeo baseada em movimento”, XXV Simpósio Brasileiro de Telecomunicações, Recife, 2007. DVD.
- [8] S. Brofferio; L. Carnimeo; D. Comunale; G. Mastronardi, “A Background Updating Algorithm for Moving Object Scenes”, In: CAPPELLINI, V. (ed.). Time-varying Image Processing and Moving Object Recognition, Amsterdam, The Netherlands: Elsevier Publishers B.V., 1997, pp. 289-296.
- [9] E. Cantu, “Multicast de vídeo digital gerado por sistema de vigilância sobre rede TCP/IP”, Projeto de pesquisa PIBITI, 2008.
- [10] C.J. Salvador; P.H. Rosa; D.S. Medeiros; F.S. Silva; M. Neto Noronha; M. Moecke, “Transmissão de seqüências de imagens através da rede celular GSM/EDGE”, II Jornada da Produção Científica da Educação Profissional e Tecnológica da Região Sul, Pelotas, 2008.
- [11] M. Moecke, “Controle de uma câmera pelo movimento detectado através de sinais de vídeo”, Projeto de pesquisa PIBITI, 2008.

AGRADECIMENTOS

Pelo suporte financeiro através de bolsa PIBITI ao CNPq e pelo espaço físico e materiais disponibilizados para a realização do projeto ao IFSC.