

Sistema de Visão Computacional para Extração Automática de Parâmetros em Ambientes Monitorados

Mateus Moraes Bueno, Maurício Edgar Stivanello, Saulo Varga, Emerson André Fedechen
Instituto Federal de Santa Catarina - IFSC
Gaspar, SC, BRASIL

Email: mateusmobueno@hotmail.com, maustiva@das.ufsc.br, saulo.vargas@ifsc.edu.br, emerson.fedechen@ifsc.edu.br

Keywords-Calibração; Detecção; Monitoramento; Segmentação; Visão Computacional.

I. INTRODUÇÃO

Os sistemas de visão computacional tem sido cada vez mais empregados para análise de fluxos de pessoas ou objetos em movimento. Através de câmeras instaladas em ambientes estratégicos, sinais de vídeo são capturados e utilizados para visualização e extração de parâmetros diversos, por meio de técnicas de processamento de imagem e visão computacional.

O desenvolvimento de sistemas de visão computacional, corrobora para a detecção de eventos em tempo real que auxiliam o ser humano na análise de ambientes monitorados. Neste sentido, o objetivo do presente trabalho é desenvolver um sistema de visão para detecção de objetos e estimação de valores para parâmetros como posição, trajetória, tempo e velocidade em ambientes monitorados.

II. METODOLOGIA EMPREGADA

Para o desenvolvimento do sistema estão sendo considerados como requisitos funcionais principais a detecção de objetos em movimento e a estimação de sua posição métrica em um ambiente monitorado. Com estas informações é possível gerar automaticamente eventos ou extrair parâmetros como velocidade, trajetória e distâncias percorridas.

As principais etapas que descrevem o processamento realizado no sistema em desenvolvimento são apresentadas na figura 1. Além da etapa inicial de preparação do sistema, podem ser observadas etapas clássicas de processamento de imagens muito comuns em sistemas de inspeção e monitoramento automatizados [1].

Para cada imagem capturada, inicialmente são detectados elementos de interesse. A detecção é realizada com base em segmentação por subtração da imagem de fundo, filtragem por operações morfológicas e agrupamento por conectividade. Os agrupamentos de pixels, detectados como objetos, são então descritos adequadamente com base em sua forma, área e coordenada de base.

Conforme ilustrado na figura 2, para obter a posição real de cada objeto detectado é necessário converter as coordenadas em pixels no sistema de coordenadas da imagem

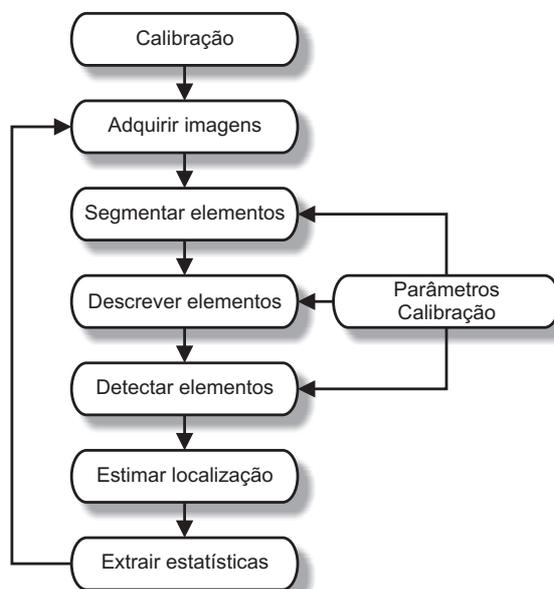


Figure 1. Fluxo do processamento de imagens

para coordenadas métricas, segundo um dado sistema de coordenadas do mundo. A conversão de coordenadas entre os sistemas de coordenadas envolvidos foi conseguida através da estimação da Transformação Projetiva que relaciona os dois planos. Os coeficientes envolvidos na transformação são estimados através de um procedimento de calibração, que ocorre na etapa de preparação do sistema. A calibração é baseada na associação de pontos com coordenadas conhecidas no mundo e na imagem [2], [3]. Com base na posição métrica dos objetos ao longo das aquisições são estimadas novas informações que podem ser utilizadas para simples registro ou para geração automática de eventos.

O sistema vem sendo desenvolvido em C++ e Matlab, fazendo uso de funções de processamento de imagens disponibilizadas pela biblioteca OpenCV.

III. RESULTADOS PRELIMINARES

Para verificação objetiva dos resultados obtidos com as funcionalidades implementadas no sistema foram avaliados

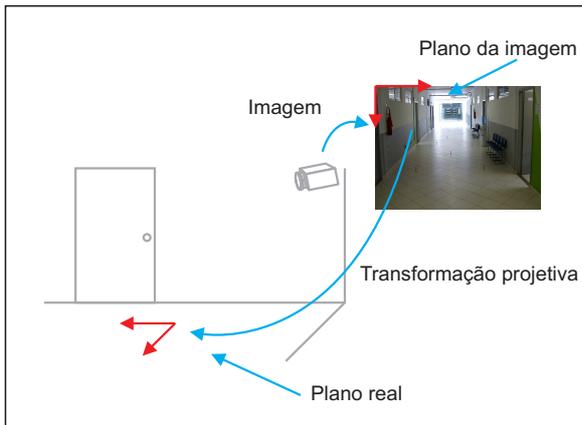


Figure 2. Transformação de coordenadas

dois cenários distintos. O primeiro, exibido na figura 3 - a, representa uma situação de monitoramento predial. Neste caso, o monitoramento é interessante para geração de eventos que indiquem a presença, trajetória ou permanência de pessoas em ambientes controlados. O segundo cenário, exibido na figura 3 - b, corresponde ao emprego do sistema para extrair informações para aplicações no esporte. Aqui, parâmetros como velocidade, distância percorrida ou posicionamento podem ser interessantes para avaliação de performance de atletas, por exemplo.

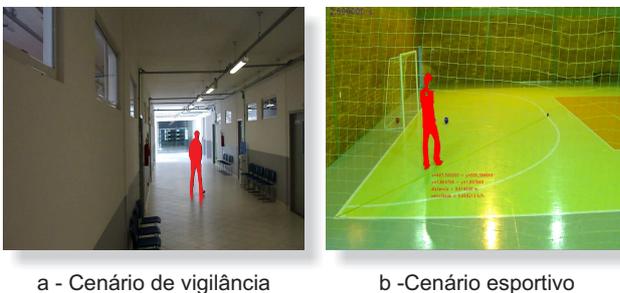


Figure 3. Cenários de aplicação

De modo geral, a detecção de pessoas ou objetos em movimento na cena observada mostrou-se efetiva para ambos cenários, apesar da simplicidade da técnica em que é baseada. Esta abordagem mostra-se adequada visto que nas situações consideradas é desejado discriminar objetos em movimento onde existe a presença de um fundo estático. Entretanto, observa-se que, ao longo do tempo, mudanças de iluminação na cena observada podem ocorrer ocasionando efeitos indesejados como sombras e reflexos. Desta forma, faz-se necessário atualizar periodicamente a imagem de referência. Além disso, é importante observar que as imagens resultantes devem ser processadas por filtros morfológicos, como erosão ou dilatação, para eliminar a presença de pequenos elementos que não representam objetos da cena, derivados de ruído gaussiano ou outros artefatos presentes nas imagens adquiridas.

Outros ensaios foram realizados com o objetivo de verificar os resultados obtidos para o caso de uso de estimação de coordenadas métricas. Para este fim, foram capturadas imagens contendo objetos cuja coordenada métrica no mundo era conhecida. Para o caso do monitoramento predial, por exemplo, foi considerado o piso do andar monitorado como sistema de coordenadas do mundo, tendo a origem convencionada como um dos cantos do mesmo. Foi considerado para estimação das coordenadas o ponto central da base de cada objeto. Para cada objeto, foi identificado o respectivo ponto na imagem capturada pela câmera. Este ponto foi então utilizado para realizar a retroprojeção no espaço, empregando a transformação estimada na calibração, obtendo-se assim a estimação do ponto 3D no mundo a partir do ponto 2D na imagem. Neste tipo de ensaio, considerando uma área de observação de até 15 metros, os pontos estimados apresentaram uma distância euclidiana média de 0,1468 m, com desvio padrão de 0,17 m. Isso significa que não há diferença significativa entre os pontos estimados e os pontos esperados, com $p = 0,98$.

Em relação ao tempo de processamento, para a etapa de detecção de objetos em movimento foi obtida uma média de 45 ms em um computador com processador Pentium D de 2.80GHz e 2GB de memória RAM. A etapa de conversão de coordenadas, por sua vez, corresponde a um produto matricial, e desta forma apresenta custo praticamente desprezível se considerarmos o tipo de aplicação.

Tendo como base a detecção e a estimação das coordenadas, outras informações têm sido avaliadas. O espaço percorrido, por exemplo, pode ser estimado pelo somatório das distâncias entre duas coordenadas consecutivas. Por sua vez, a velocidade pode ser estimada pela razão entre o espaço percorrido e a informação de tempo, obtido com base na taxa de aquisição. O resultado na estimação destas, e outras informações, tem se mostrado efetivo.

IV. CONSIDERAÇÕES FINAIS

O objetivo do trabalho descrito é a implementação de um sistema de monitoramento automático por meio de técnicas de visão computacional. Considera-se que os resultados de detecção e estimação alcançados até o momento são satisfatórios, e que a precisão das informações calculadas é suficiente para uma vasta gama de aplicações de monitoramento de ambientes.

Como trabalhos futuros pode-se citar a implementação de uma técnica robusta para rastreamento de objetos e a utilização de um modelo que permita tratar de modo global as informações provenientes de múltiplas câmeras.

REFERENCES

- [1] R. C. Gonzalez and R. E. Woods, *Processamento de Imagens Digitais*. Edgard Blucher, 2000.
- [2] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.