

Relatório Projeto 3 de Introdução ao Processamento de Imagens (IPI) - 2024/1

Pedro Rodrigues Diógenes Macedo
Departamento de Ciências da Computação,
Brasília, Brasil
prdm.002@gmail.com

Cássio Vinícius Teixeira Borges
Departamento de Ciências da Computação,
Brasília, Brasil
cassiocvtb2003@gmail.com

Abstract—Este relatório discute os estudos e experimentos realizados no projeto 3 da disciplina de Introdução ao Processamento de Imagens (IPI) que envolve a tentativa de implementar a solução do artigo *A Method for Detecting Pedestrians in Video Surveillance Scenes*. Os experimentos envolvem a extração de objetos em movimento por diferença de multi-quadros, algoritmo de iluminação ambiente adaptativa para extração de características da pele e por fim a transformação de Hough melhorada para extração de cabeça humana. Por fim, os resultados de cada experimento são apresentados e desenvolvidos.

I. INTRODUÇÃO

O rápido avanço das tecnologias de processamento digital de imagens e comunicação em redes acelerou o desenvolvimento de sistemas inteligentes de vigilância por vídeo. Esses sistemas são importantes pois incluem controle de acesso, identificação de pessoas e análise de comportamentos.

A detecção de pedestres em cenários urbanos apresenta diversos desafios devido a fatores dinâmicos, como variações de iluminação, locais fechados e cenários complexos. A fim de obter algum resultado positivo nessa área, várias técnicas têm sido exploradas, como a subtração de fundo, histogramas de gradiente e modelos baseados em vistas. No entanto, muitos desses métodos sofrem com a ocorrência de falsos positivos, especialmente quando se baseiam em uma única abordagem de detecção.

Neste relatório, buscamos aplicar o método proposto por Zhang (ZHANG et al., 2012) para a detecção de pedestres, utilizando uma combinação de informações de movimento, detecção da cor da pele humana e ajustes adaptativos de iluminação ambiente. Além disso, faremos uso da técnica de diferenciação de múltiplos quadros e a transformada de Hough para melhorar a detecção de pedestres, mesmo sob condições variáveis de iluminação. Nossa objetivo é verificar os métodos e realizar sua aplicação utilizando o *Python* e a biblioteca *OpenCv*.

II. METODOLOGIA

Nesta seção, descrevem-se os métodos utilizados sendo eles a extração de objetos em movimento por diferença de multi-quadros, algoritmo de iluminação ambiente adaptativa para extração de características da pele e transformação de Hough melhorada para extração de cabeça humana.

A. Extração de objetos em movimento por diferença de multi-quadros

A extração de objetos em movimento é feita pela comparação da intensidade de pixels entre quadros consecutivos de uma sequência de vídeo. A diferença entre dois quadros pode ser calculada pela seguinte equação:

$$D(x, y) = \begin{cases} 1 & \text{se } |F_{i+1}(x, y) - F_i(x, y)| > T \\ 0 & \text{caso contrário} \end{cases}$$

onde $F_i(x, y)$ é a intensidade do pixel (x, y) no quadro i , e T é um limiar pré-definido. A diferença absoluta entre os quadros é usada para identificar áreas de movimento, que são então binarizadas e refinadas através de operações de dilatação e erosão

B. Algoritmo de iluminação ambiente adaptativa para extração de características da pele

A cor da pele humana é uma característica interessante na detecção de pedestres, especialmente em vigilância visual, onde a resolução da imagem pode ser limitada e a iluminação pode variar consideravelmente.

Para lidar com as mudanças na iluminação ambiente, o algoritmo adapta os valores da cor da pele humana com base na média de iluminação da imagem, usando o modelo HSI (matiz, saturação e intensidade). A equação para calcular esses valores é dada por:

$$H = \left(\frac{\mu - 0.15}{10} \pm 0.012 \right), \quad S = \left(\frac{\mu + 0.1}{10} \pm 0.070 \right), \\ I = (1.12 \cdot \mu \pm 0.060)$$

onde μ é a média da luminosidade da imagem. A imagem é então convertida para o espaço de cor HSV, e as locais que correspondem à pele são extraídas aplicando máscaras nos canais de matiz, saturação e intensidade. Esses valores são ajustados dinamicamente com base na iluminação.

C. Transformação de Hough melhorada para extração de cabeça humana

A Transformada de Hough é utilizada para detectar o formato da cabeça humana, assumindo que a cabeça tenha uma forma aproximadamente circular. O raio estimado da cabeça é

baseado na proporção entre a cabeça e o corpo humano, que é dada por:

$$r = \frac{h \times 65}{1000}$$

onde h é a altura do corpo detectado. A Transformada de Hough é então aplicada para detectar círculos com raios próximos a r . A equação da transformada de Hough é:

$$(x - a)^2 + (y - b)^2 = r^2$$

III. RESULTADOS

A implementação dos procedimentos descritos não apresentou os resultados esperados. Conforme descrito, cada objeto detectado entre dois frames deveria ser classificado como pedestre ou não, utilizando os procedimentos de detecção de características de pele e de detecção de cabeças¹. No entanto, esses dois processos foram implementados de maneira que poucos objetos foram classificados corretamente como pedestres, o que não refletiu a realidade dos frames analisados. Como o artigo utilizado como base não especifica detalhadamente como cada procedimento deve ser realizado, acreditamos que a forma como implementamos esses passos pode não ter sido a mais apropriada, resultando na ineficiência do programa para essas etapas.



Figure 1. Resultado da detecção de cabeças

Fazendo uso da função *HoughCircles* do *OpenCv* em alguns testes conseguimos detectar algumas áreas do corpo em movimento das pessoas nas gravações, a cabeça sendo uma delas. Contudo, não conseguimos dar sequência na implementação como discutido no artigo (ZHANG et al., 2012).

Por outro lado, o procedimento de detecção de objetos em movimento apresentou resultados mais satisfatórios. Em testes com vídeos contendo poucos objetos em movimento e pedestres caminhando em linha reta, conseguimos resultados que se aproximam do comportamento esperado. Na Figura 2,

por exemplo, a detecção foi realizada conforme o esperado. Já na Figura 3, obtivemos boas detecções quando um pedestre caminha sozinho; no entanto, quando pedestres estão próximos uns dos outros, o programa os considera como um único objeto. Na Figura 4, observa-se um caso em que o movimento da água no lago ao fundo interfere na detecção do pedestre, causando dificuldades no processo de detecção.

Esses resultados foram gerados com os procedimentos de classificação de pedestres desativados, pelos motivos explicados anteriormente.



Figure 2. Resultado com apenas um pedestre.



Figure 3. Resultado com mais pedestres em movimento



Figure 4. Resultado com um pedestre com movimento no fundo.

IV. CONCLUSÃO

O projeto ofereceu uma oportunidade valiosa para aplicar os conhecimentos adquiridos durante a disciplina na tentativa de replicar uma solução científica descrita em um artigo. Embora nem todos os resultados esperados tenham sido alcançados em relação a cada tópico da metodologia, conseguimos obter resultados significativos na detecção de pessoas e grupos de pessoas em vídeos capturados em locais abertos, com diferentes condições de iluminação. Utilizamos uma combinação de informações de movimento, e formato do corpo humano para enfrentar os desafios da detecção de pedestres em cenas de vídeo, conforme discutido no artigo base. Contudo, vale ressaltar a dificuldade que tivemos em encontrar vídeos de câmera de vigilância, isso nos fez trabalhar com vídeos similares não coletados por câmeras de vigilância.

REFERENCES

- ZHANG, Xin et al. A Method for Detecting Pedestrians in Video Surveillance Scenes. *IEEE*, 2012. Available from: <https://ieeexplore.ieee.org/document/6223446>.