

Sistema para Detecção da Trajetória do Movimento de Golf Swing a Partir de Vídeos Gravados de um Smartphone

Jordan de Oliveira Kau Lyrio¹, Karin Satie Komati², Daniel Ribeiro Trindade¹

¹Coordenadoria de Informática (CIn)

²Programa de Pós-Graduação em Computação Aplicada (PPComp)
Campus Serra do Instituto Federal do Espírito Santo (Ifes)
Serra - ES - Brasil

jordankl93@gmail.com, {kkomati,danielrt}@ifes.edu.br

Abstract. *In golf, the movements performed by the player are very important. Among these, one that stands out is the golf swing movement. There are several applications on the market that help players to improve their golf swing. However, these applications are generally not free or even require the user to wear additional equipment or complex camera configurations to track their movement. In this work we present an algorithm to track the a golfer's golf swing movement. The system uses image processing techniques to detect and display the trajectory of the golf club. Videos recorded with a smartphone are used as input for the system. In order to assess the system's effectiveness, the detected trajectory is compared with the real trajectory, wich is defined manually. For the experiments, seven videos with different swing styles and types of background scenarios were used. The results obtained were satisfactory for most of the cases.*

Resumo. *No golfe, os movimentos realizados pelo jogador são muito importantes. Dentre esses, um que se destaca é o movimento de golf swing. Existem diversas aplicações no mercado que auxiliam os jogadores a melhorarem seu movimento. Entretanto, essas aplicações são pagas ou exigem que o golfista utilize equipamentos acessórios ou configurações complexas de câmera para rastrear seu movimento. Neste trabalho foi desenvolvido um sistema de rastreamento do movimento de golf swing de um golfista. O sistema exibe a trajetória do taco a partir de vídeos gravados da câmera de um smartphone, usando técnicas de processamento de imagens. Para aferir a eficácia do sistema, é comparada a trajetória detectada pelo algoritmo com a trajetória real, essa última definida manualmente. Para os experimentos, foram utilizados sete vídeos de diferentes estilos de tacadas e com diferentes tipos de cenários de fundo. Os resultados obtidos foram satisfatórios para a maioria dos casos.*

1. Introdução

O golfe é um esporte que vem crescendo muito nas últimas décadas, e já é praticado por mais de 60 milhões de pessoas ao redor do mundo [Wheeler and Nauright 2006]. Um dos objetivos do jogo é acertar os buracos espalhados em um campo no menor número de tentativas possíveis ou, dependendo da modalidade, acertar o maior número de buracos em uma seção. A partida se inicia em uma área chamada de *tee*, onde é realizada a tacada

inicial de um percurso. Neste momento, o golfista costuma realizar a tacada drive, que é executada para lançar a bola em longa distância. A área de chegada, chamada de *green*, é onde fica localizado o buraco. Nesta área, os jogadores costumam realizar a jogada de *putting*, que é a tacada feita para colocar a bola no buraco. Para cada tacada realizada, o movimento realizado pelo corpo é fundamental para conseguir uma jogada com alto nível de eficácia [Riveiro et al. 2015].

Dentre as diferentes técnicas de execução de uma tacada, destacam-se os movimentos de *golf putting* e *golf swing*. No *golf putting*, o movimento realizado pelo corpo reproduz uma forma pendular que visa acertar a bola com maior precisão possível a fim de acertar o buraco. No *golf swing*, o movimento realizado pelo corpo do golfista exige uma atividade coordenada dos músculos para geração de energia que será transferida para bola a fim de realizar uma tacada de longa distância [McHardy and Pollard 2005]. Por envolver movimentos complexos e que exigem grande trabalho de coordenação dos principais membros do corpo, os jogadores amadores costumam gastar grande parte do seu tempo tentando aperfeiçoar a realização do *swing* e do *putting*.

Existem diversos tipos de aplicações que analisam de diferentes modos o *swing* de um golfista. Alguns desses sistemas necessitam de roupas especiais para rastrear o movimento corporal do jogador ou de um taco com sensores ligados a ele. No entanto, a necessidade de utilizar dispositivos conectados ao corpo ou ao taco de golfe, em muitos casos, acaba tornando a experiência desagradável, pois tais sensores podem tornar o taco mais pesado ou causar desconforto no momento de realizar o *golf swing*. Além disso, algumas aplicações não são tão acessíveis, por serem muito caras ou exigirem que o golfista já tenha alguma experiência no golfe [Chun et al. 2014].

Este artigo descreve um sistema para detectar o movimento de *swing* de um golfista utilizando como entrada o vídeo da jogada. Isso permite que o usuário utilize *smartphones* para a geração do vídeo. Como resultado o sistema exibe a trajetória do movimento de *swing*. Para alcançar esse resultado foram utilizadas técnicas de processamento de imagens para realizar a detecção do taco e acompanhar a trajetória da ponta do taco nos vários quadros do vídeo. Para aferir o funcionamento do sistema foi calculada a diferença entre as trajetórias reais e as detectadas pelo sistema.

2. Trabalhos Correlatos

No trabalho de [Urtasun et al. 2005] foi proposta uma abordagem que incorpora modelos dinâmicos no processo de rastreamento do corpo humano para produzir uma representação completa em 3D a partir de sequências monoculares. A abordagem utiliza modelos de movimentos específicos do *golf swing* evitando a complexidade computacional de algoritmos multi-hipóteses.

No trabalho de [Karlga and Hwang 2006] foram analisados movimentos do corpo humano com base em um modelo 3D obtido a partir de sequências de vídeo de uma câmera única durante o *golf swing*. São utilizadas as informações bidimensionais do corpo humano, extraídas através de técnicas de segmentação de objetos, e obtenção do modelo 3D usando um algoritmo iterativo de ajuste 3D e Redes Dinâmicas Bayesianas (*Dynamic Bayesian Networks*). A partir das informações de movimentos obtidas do modelo 3D, comparou-se o *golf swing* de diferentes jogadores, independente da

variabilidade da aparência humana, da perspectiva da câmera, da complexidade do movimento humano e diversos outros fatores de complexidade. De acordo com [Karlga and Hwang 2006], os resultados obtidos demonstraram que o sistema é capaz de captar a estrutura do corpo humano tridimensionalmente nos movimentos de *golf swing*. Em [Karlga and Hwang 2007] o algoritmo é aplicado com o diferencial de se analisar as principais partes do corpo humano separadamente e, assim, obter mais precisão na representação do modelo 3D.

No trabalho de [Couceiro et al. 2013] foi apresentada uma metodologia para detecção visual e estimação de parâmetros na análise da variabilidade no desempenho de *golf putting* utilizando uma câmera digital para realizar o rastreamento do movimento. A detecção da posição do taco horizontalmente é obtida utilizando técnicas de visão computacional e de um algoritmo de estimação divididos em duas etapas. Na primeira parte algumas técnicas de estimação não-lineares com desempenho diferentes são analisadas para extração de um modelo senoidal. Na segunda parte são realizados experimentos com jogadores de golfe experientes e, com base nos resultados da primeira parte, é utilizada a técnica *Darwinian Particle Swarm Optimization* (DPSO) para obter uma caracterização da técnica de *putting* de cada jogador. Os resultados obtidos por [Couceiro et al. 2013] mostram que a aplicação proposta é bem funcional e permite recuperar uma série de informações simultâneas sobre o *golf putting* de um jogador.

No trabalho de [Fung et al. 2014] foi desenvolvido um sistema rastreador de movimento humano sem uso de marcadores que é capaz de monitorar as principais partes do corpo de um atleta em um vídeo de transmissão esportiva. Para o sistema, foi proposto um método de rastreamento híbrido que consiste na combinação de três algoritmos para rastrear a cabeça, o corpo, as mãos, os ombros, os joelhos e os pés de um golfista na realização do movimento de *golf swing*. Os três algoritmos utilizados foram fluxo óptico piramidal de Lucas-Kanade (LK), correspondência de modelos baseados em correlação normalizada e subtração de fundo. Com os resultados do rastreamento, uma representação do corpo humano em forma de esqueleto é apresentada para imitar a pose exata do jogador na execução do movimento.

No trabalho de [Riveiro et al. 2015] foi realizada uma pesquisa de campo com instrutores de golfe a fim de se levantar quais recursos um sistema de treinamento baseado em computador deve ter para auxiliar o ensino do *golf swing*. Para orientar o desenvolvimento destes sistemas diversas questões foram levantadas, entre elas qual a característica de um bom movimento de *swing*, como são realizadas as sessões de treinamento e os desafios enfrentados pelos instrutores. Os resultados obtidos por [Riveiro et al. 2015] mostram que um sistema de treinamento de *golf swing* baseado em computador deve possuir características de comparação entre o *swing* de um jogador habilidoso e um jogador novato.

No trabalho de [Gehrig et al. 2003] foi apresentada uma metodologia de rastreamento visual que utiliza um modelo global de movimento. O objetivo do trabalho é recuperar a trajetória da ponta do taco de golfe em espaço-temporal 2D a partir de vídeo de *golf swing*, de modo que informações como velocidade do taco, orientação e aceleração também possam ser extraídas. O método desenvolvido é totalmente automatizado, dispensando a interferência de um usuário ou uso de equipamentos e sensores para realizar a detecção. O sistema também pode ser utilizado em cenários com fundos bastante

complexos. O algoritmo se adapta de forma robusta ao modelo de *swing* através das hipóteses de localização do taco de golfe a partir de quadros únicos e também dispensa o uso de equipamentos de detecção de movimentos.

3. O Sistema Proposto

A Figura 1 apresenta a arquitetura geral do sistema. Para cada quadro do vídeo, há quatro passos sequenciais: a detecção de movimento, a detecção de possíveis tacos, a detecção do taco e a exibição do taco no quadro. Cada um dos passos será detalhado nas subseções a seguir. O sistema foi desenvolvido usando a linguagem Java e as funções disponíveis na biblioteca OpenCV para operações de processamento digital de imagem (PDI) [Gonzalez and Woods 2002].

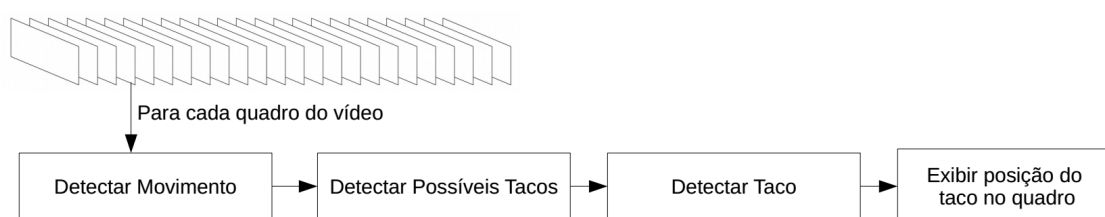


Figura 1. Arquitetura geral proposta.

3.1. Detecção do movimento

Para rastrear o movimento do taco foram aplicadas técnicas de processamento de imagens nos quadros do vídeo. A Figura 2 apresenta o resultado após cada operação de PDI realizada.

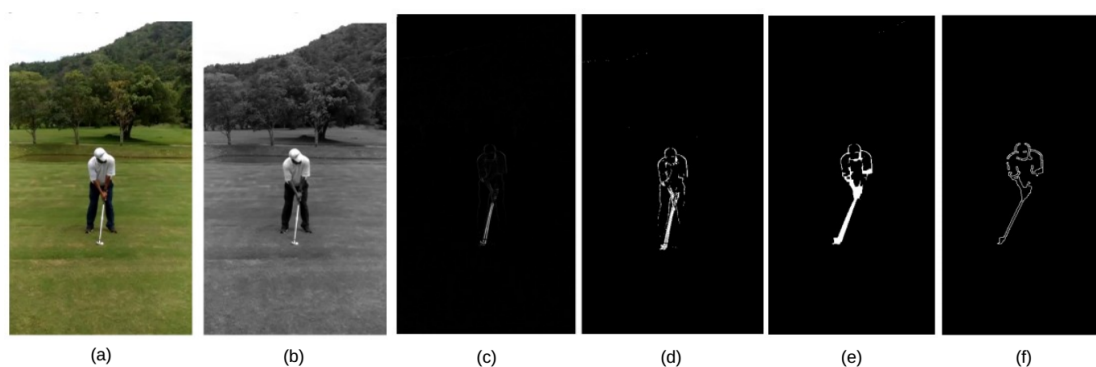


Figura 2. Resultados de cada processamento de imagens na detecção de movimentos.

Para um dado instante do vídeo, são utilizados 3 quadros: o anterior Q_{n-1} , o atual Q_n e o próximo Q_{n+1} . Esses quadros são convertidos para escala de cinza (Figura 2(b)). São gerados então dois novos quadros: D_1 , resultante da diferença entre o próximo quadro e o atual ($Q_{n+1} - Q_n$); e D_2 , resultante da diferença entre o próximo quadro e o atual ($Q_n - Q_{n-1}$). A Figura 2(c) mostra um desses quadros.

Após a geração de D_1 e D_2 , é realizada uma operação de limiarização (em inglês *threshold*) nesses quadros. Esta operação transforma uma imagem em escala de cinza em

uma imagem binária (*black and white*), onde os objetos com valor maior que o limiar são transformados em pixels de cor branca e os com menores valores, em pixels de cor preta. Após testes preliminares, chegou-se a valor de limiar 15. Em seguida, foi aplicada a operação de fechamento morfológico (*morphological closing*) em D_1 e D_2 . O objetivo do fechamento morfológico é eliminar pequenas falhas nas regiões de movimento (preenchendo pequenos furos dentro de objetos) e suavizar as bordas encontradas. A Figura 2(d) é o resultado da aplicação das operações de limiar e fechamento.

O passo seguinte é a geração de uma nova imagem R , resultante da interseção entre D_1 e D_2 . Para isso aplicou-se o operador lógico AND. A imagem R pode ser vista na Figura 2(e). Ao final, foi aplicado em R o algoritmo Canny [Canny 1986] para extração das bordas (Figura 2(f)).

3.2. Detecção dos possíveis tacos

Após o tratamento da imagem para identificarmos os objetos em movimento no vídeo, o próximo passo a ser feito é a identificação de quais desses objetos são possíveis tacos de golfe. É esperado que o usuário indique os pontos onde inicialmente se encontram a cabeça do taco e as mãos do golfista. Isso serve como ponto de partida para o algoritmo de detecção do taco. Como é um dado fornecido pelo próprio usuário, então é razoável afirmar que o segmento de reta formado por esses pontos tem o tamanho do taco de golfe na imagem. O usuário é capaz de escolher esses pontos com o mouse, caso esteja usando um computador, e com um toque de tela no caso de uso de um *smartphone*.

No próximos quadros, foi utilizada a transformada de Hough [Duda and Hart 1972] para extrair segmentos de retas presentes na imagem. Assume-se que cada segmento de reta encontrado é um possível taco. Cada segmento de reta é formado pelos dois pontos de extremidade do possível taco. A partir desses pontos são calculados o coeficiente angular da reta e o tamanho do segmento encontrado.



Figura 3. Imagem que apresenta o raio de busca em círculo verde.

Com as informações iniciais fornecidas pelo usuário é possível estabelecer uma área onde é mais provável que o taco esteja. Considerando que o movimento de *swing* é similar a um arco, então foi definida uma área circular com centro no ponto inicial da mão do jogador e raio igual a 2,5 o comprimento do taco (circulo em verde na Figura 3). Segmentos de reta que estejam localizados fora dessa área são desconsiderados (segmentos em amarelo, na parte superior da Figura 3). Também são desconsiderados segmentos cujo tamanho seja muito menor ou maior em relação ao segmento de reta fornecido pelo usuário.

Após a aplicação do algoritmo Canny, utilizado ainda na etapa de detecção de movimento, verifica-se que duas retas paralelas são formadas no contorno do taco de golfe. A Figura 4 apresenta a mesma Figura 2(f), porém amplificada apenas na região do golfista. Para encontrar as retas paralelas utiliza-se o coeficiente angular de cada reta para realizar a comparação entre elas. Todas as retas são consideradas paralelas se a diferença de coeficiente angular for menor que 0,5, valor empírico. Cada par de retas paralelas forma um quadrilátero. O segmento de reta candidato a taco é segmento diagonal desse quadrilátero.

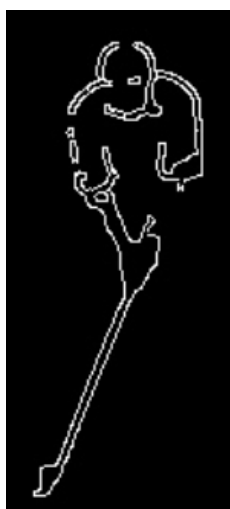


Figura 4. Amplificação do resultado da aplicação do filtro de Canny.

3.3. Detecção do taco e exibição do taco no quadro

Esta é a etapa final do algoritmo e tem como objetivo determinar qual é o segmento correto entre os candidatos fornecidos pelo passo anterior. Uma vez escolhido esse segmento também é necessário definir qual, dentre os dois pontos que compõem o segmento escolhido, representa a cabeça do taco.

Dentre os segmentos de reta candidatos, verifica-se qual deles é o mais próximo ao último segmento de reta escolhido como taco em quadros anteriores. Isso é feito calculando-se a distância euclidiana [Pedott and Fogliatto 2013] entre os segmentos candidatos e o segmento escolhido anteriormente. Uma vez determinado o segmento que representa o taco, define-se qual de seus pontos de extremidade corresponde à mão do jogador: o ponto do taco atual mais distante do ponto do taco anterior referente à mão do jogador é definido como a cabeça do taco atual.

Para visualizar a trajetória do movimento de *swing*, o novo ponto de cabeça do taco encontrado é ligado com uma linha preta ao ponto referente à cabeça do taco anterior, gerando um traçado da trajetória do taco (Figura 3).

4. Experimentos, Resultados e Discussão

Nesta seção apresentam-se os resultados obtidos pela aplicação, bem como a discussão dos resultados. Foram utilizados 7 (sete) vídeos coletados pelos próprios autores em um campo oficial de golfe. Os vídeos foram feitos a partir da câmera de um *smartphone* sem a utilização de um tripé ou qualquer equipamento para estabilizar os vídeos. Para se obter uma taxa maior de quadros, os vídeos foram capturados com a função *slow motion* ativada. Os jogadores que foram filmados durante a execução dos movimentos são profissionais do esporte. Cada um dos vídeos tem resolução de 960 x 720 quadros. Os vídeos gravados e utilizados para o experimento podem ser encontrados no site <https://github.com/jordankl93/GolfSwingAnalyser/tree/master/videos>.

Para a análise quantitativa, avaliou-se a precisão do algoritmo de detecção. A aplicação grava a posição do taco em cada quadro quadro, armazenando as coordenadas da trajetória da ponta do taco. O *benchmark* foi feito manualmente em cada quadro. Para cada par de coordenadas, do ponto detectado pelo algoritmo e o manual, é calculada a distância euclidiana entre eles. As distâncias encontradas são denominadas de Erro. Os valores de Erro, por sua vez, são exibidos no término do processamento e, posteriormente, são utilizados para calcular a Média de Erro da trajetória detectada pelo algoritmo.

Além da Média de Erro calculada, também foi feito o cálculo do Desvio Padrão dos Erros para cada vídeo utilizado [Pedott and Fogliatto 2013]. O Desvio Padrão é utilizado para verificar a uniformidade de um conjunto de dados, ou seja, quanto mais próximo o resultado for de 0 (zero) mais uniformes são os dados. Para o nosso caso, a medida do desvio padrão mostra o quão próximo a trajetória detectada está da trajetória real. Essas medidas são importantes para mostrarem o grau de eficácia do algoritmo em diferentes cenários.

Os resultados quantitativos são apresentados na Tabela 1 e mostram a diferença nos resultados das detecções nos diferentes cenários e estilos de *golf swing* realizados pelos golfistas. A primeira coluna apresenta a quantidade total de quadros do vídeo. A segunda coluna contém o número de quadros onde houve detecção. Esses valores são menores em relação ao total de quadros do vídeo pois nem sempre ocorre a detecção do taco. A terceira coluna apresenta a média dos erros as posições detectadas pelo sistema e as posições definidas manualmente. A quarta coluna apresenta os valores do desvio padrão do erros calculados.

Neste trabalho foi utilizado como métrica o pixel. No Vídeo 1, a média de distância de erro foi de 9,172 pixels. Já o desvio padrão foi de 4,69. Este vídeo possui o segundo melhor resultado de detecção quando avaliamos as medidas utilizadas. Na Figura 5 observa-se a diferença da trajetória manual (cor azul) e a trajetória detectada pelo algoritmo (cor preta) do Vídeo 1.

A Figura 6 apresenta os resultados dos vídeos de 2 à 7, na ordem de (a) a (f). O Vídeo 2 (Figura 6(a)) é o que possui o melhor resultado avaliando as medidas calculadas. Observa-se que a trajetória do algoritmo ficou bem próxima do *benchmark*, mas ao mesmo

Tabela 1. Resultado quantitativo dos vídeos de 1 a 7.

	Quantidade total de quadros	Quantidade total de quadros com detecção	Média de Distância de Erro (em pixels)	Desvio Padrão
Vídeo 1	103	82	9,172	4,690
Vídeo 2	81	66	6,319	3,154
Vídeo 3	104	26	10,242	7,062
Vídeo 4	90	33	11,376	6,959
Vídeo 5	84	49	40,348	81,520
Vídeo 6	96	47	19,642	14,724
Vídeo 7	95	44	13,202	16,928

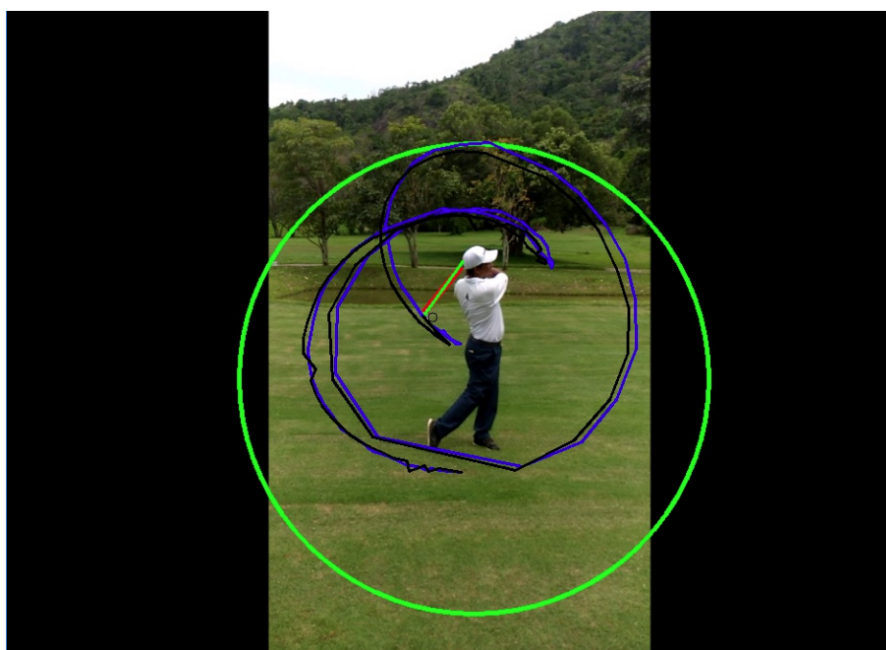


Figura 5. Resultado qualitativo do Vídeo 1.

tempo é o movimento curto, de menor extensão e o jogador não realizou movimentos acentuados.

O Vídeo 5 apresentou os piores resultados, com média de distância de erro calculada de 40,348 pixels e desvio padrão de 81,520. Na Figura 6(d) observa-se que a trajetória detectada pelo algoritmo, exibida em preto, realmente é muito diferente do *benchmark*. Uma questão no vídeo 5 é a espessura do taco, que era mais fino [Brito et al. 2012]. Um golfista pode carregar até 14 tacos em um jogo, tendo variações em seu tipo, material, espessura, dentre outros. Como o algoritmo assume que um taco de golfe necessariamente deve possuir um par de retas paralelas e neste vídeo o taco de golfe possui uma espessura muito fina, o algoritmo identifica apenas uma reta no taco, não realizando a detecção corretamente.

5. Considerações Finais

Neste trabalho foi desenvolvido um sistema para rastrear o movimento de *golf swing* de um golfista utilizando vídeos gravados de uma câmera de celular. O sistema realiza uma

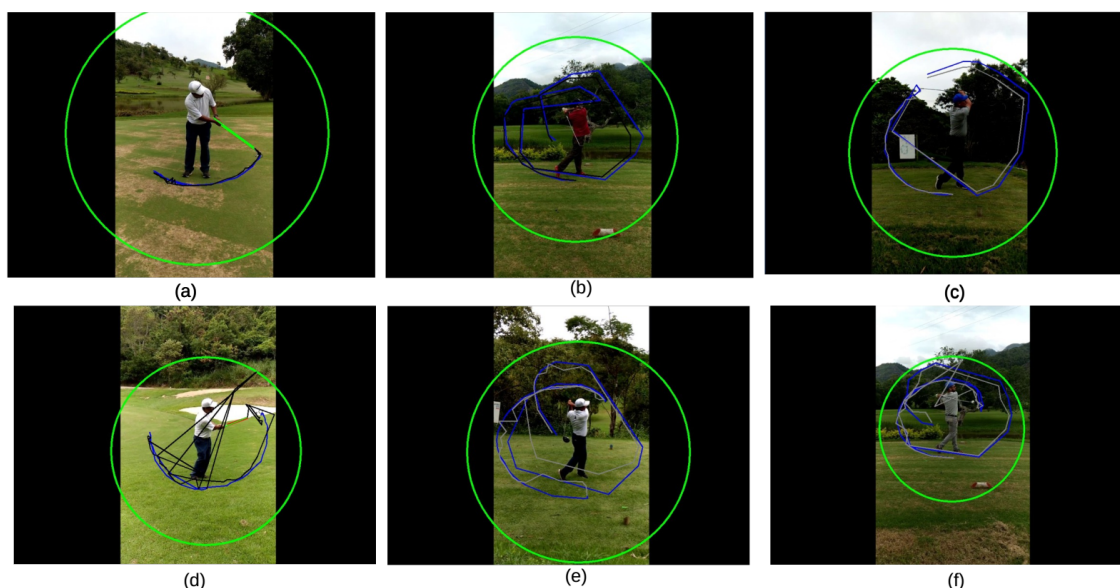


Figura 6. Resultado qualitativo dos vídeos de 2 a 7.

análise dos objetos em movimento no vídeo e utiliza características presentes no *golf swing* para identificar o taco e rastrear a trajetória do movimento.

Para que fosse possível rastrear o trajeto do taco no movimento de *swing* foi necessário realizar a detecção dos objetos em movimento no vídeo usando técnicas clássicas de PDI e conceitos de geometria para extrair as características do taco e do movimento de *swing* realizado. O resultado obtido foi exibido em tela marcando a trajetória realizada pela ponta do taco durante todo o movimento.

Ainda que os objetivos apresentados tenham sido alcançados, melhorias podem ser feitas em futuros trabalhos, tais como melhorar as escolhas dos parâmetros de detecção, extrair informações referentes a velocidade da ponta do taco no momento do *swing*, criar método para suavizar a trajetória detectada, eliminando assim pontos que tenham ficado muito longe da trajetória. Também é possível usar técnicas mais recentes, do estado da arte de aprendizado de máquina, automatizando a detecção de trajetória. É necessário que sejam realizados mais experimentos, e mais experimentos em tempo real. Além disso, seria interessante encapsular o método desenvolvido em uma biblioteca, para que ele possa ser reutilizado por aplicações finais.

Referências

- Brito, A. P., Henriques-Neto, D., Macedo, A., and Ferreira, T. (2012). Critérios de utilização dos tacos entre jogadores sêniores e não sêniores na modalidade de golfe. *Revista Digital EFDeportes.com*, 15(166).
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698.
- Chun, S., Kang, D., Choi, H.-R., Park, A., Lee, K.-K., and Kim, J. (2014). A sensor-aided self coaching model for uncocking improvement in golf swing. *Multimedia tools and applications*, 72(1):253–279.

- Couceiro, M. S., Portugal, D., Gonçalves, N., Rocha, R., Luz, J. M. A., Figueiredo, C. M., and Dias, G. (aug. 2013). A methodology for detection and estimation in the analysis of golf putting. *Pattern Analysis and Applications*, 16(3):459–474.
- Duda, R. and Hart, P. (1972). Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15.
- Fung, S. K., Sundaraj, K., Ahamed, N. U., Kiang, L. C., Nadarajah, S., Sahayadhas, A., Ali, M. A., Islam, M. A., and Palaniappan, R. (apr. 2014). Hybrid markerless tracking of complex articulated motion in golf swings. *Journal of bodywork and movement therapies*, 18(2):220–227.
- Gehrig, N., Lepetit, V., and Fua, P. (2003). Golf club visual tracking for enhanced swing analysis tools. In *Proceedings...*, pages 1–10. BRITISH MACHINE VISION CONFERENCE.
- Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Prentice-Hall, Upper Saddle River, NJ, 2 edition.
- Karliga, I. and Hwang, J.-N. (2006). Analyzing human body 3-d motion of golf swing from single-camera video sequences. In *Proceedings...*, volume 5, pages v493–v496. IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE.
- Karliga, I. and Hwang, J.-N. (2007). Extraction and integration of human body parts for 3-d motion analysis of golf swing from single-camera video sequences. In *Proceedings...*, pages 3960–3963. IEEE International Symposium on Circuits and Systems, IEEE.
- McHardy, A. and Pollard, H. (2005). Muscle activity during the golf swing. *British journal of sports medicine*, 39(11):799–804.
- Pedott, A. H. and Fogliatto, F. S. (2013). Estudos de repetitividade e reprodutividade para dados funcionais. *Production*, 23(3):548–560.
- Riveiro, M., Dahlbom, A., König, R., Johansson, U., and Brattberg, P. (2015). Supporting golf coaching and swing instruction with computer-based training systems. In *International Conference on Learning and Collaboration Technologies*, pages 279–290. Springer.
- Urtasun, R., Fleet, D. J., and Fua, P. (2005). Monocular 3d tracking of the golf swing. In *Proceedings...*, volume 2, pages 932–938. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE.
- Wheeler, K. and Nauright, J. (2006). A global perspective on the environmental impact of golf. *Sport in society*, 9(3):427–443.