
ANIMAÇÃO GRÁFICA DA MARCHA HUMANA A PARTIR DE DADOS DO KINECT

Edmo de Oliveira Leite

Gilda Aparecida de Assis

Glauco Ferreira Gazel Yared

Universidade Federal de Ouro Preto (Ufop)

Resumo

A análise da marcha humana a partir de dados biométricos tem aplicações em áreas como segurança, robótica bioinspirada e saúde. Sensores de movimento de baixo custo, como o Kinect, têm permitido a aquisição de dados biométricos da marcha em ambientes terrestres. Entretanto, esses equipamentos têm limitações que podem impactar na qualidade dos dados. Nesse cenário, diferentes técnicas de processamento de sinais podem ser aplicadas para reduzir o ruído. A visualização desses dados, originais ou processados, muitas vezes é realizada na forma de gráficos, tendo utilidade limitada para profissionais não experientes na análise de sinais. Nesse sentido, a visualização dos dados da marcha em um modelo tridimensional pode contribuir para melhorar a decisão dos profissionais, principalmente da saúde. Este trabalho tem como objetivo realizar a animação da marcha humana em um modelo tridimensional, a partir dos dados coletados pelo sensor Kinect 2.0. Para reduzir o ruído dos dados, foi realizado um pré-processamento com filtros de média móvel e Butterworth. Foram elaborados vídeos das animações conforme as vistas isométrica e lateral, que foram incorporados em um questionário *on-line* e avaliados em uma pesquisa de campo sobre artificialidade/naturalidade da animação, utilizando-se a técnica de pontuação média de opinião (*mean opinion score* [MOS]). Um total de 22 participantes, estudantes de computação, respondeu ao questionário *on-line*. A análise de variância simples (*analysis of variance* [Anova]) *one way* mostrou que os vídeos a partir das vistas isométrica e lateral processados com filtro de média móvel (janela = 15 e repetições = 3) que obtiveram

maiores valores da métrica MOS foram avaliados como significativamente mais naturais do que outros vídeos, processados ou não.

Palavras-chave: Animação. Kinect. Marcha humana.

1 INTRODUÇÃO

Estão disponíveis no mercado diversos equipamentos para realizar o procedimento de captura de movimento, desde equipamentos mais custosos e com maior acurácia, como sistemas de câmeras da Vicon Motion Systems, equipados com luz infravermelha estroboscópica e marcadores atrelados ao indivíduo ou objeto a ser medido, os quais utilizam a luz refletida pelos marcadores para calcular sua posição espacial (MIYAWAKI; SATO; KIZAWA, 2017) até equipamentos mais acessíveis, como o sensor de movimentos Kinect, uma ferramenta da Microsoft utilizada em alguns jogos do Xbox 360 (BORENSTEIN, 2012).

Os sensores de movimento de baixo custo têm permitido a aquisição de dados biométricos da marcha humana, explorada em diversas áreas, como a autenticação e a identificação do indivíduo pela caminhada, projeto de estruturas robóticas, bem como estudo e tratamento de patologias relacionadas à locomoção humana. Entretanto, a aquisição desses dados por meio desses equipamentos tem limitações que podem impactar a qualidade das informações adquiridas. Nesse contexto, diferentes técnicas de processamento de sinais podem melhorar a qualidade dos dados (TAKEDA *et al.*, 2014), como a utilização de filtro de média móvel ou filtro de Butterworth para reduzir ruídos nos sinais, atenuando sua saída (SILVA *et al.*, 2018).

A marcha humana é uma ação realizada pelo corpo humano que envolve todo o corpo, a partir de uma atividade no sistema nervoso central, envolvendo os músculos, os ossos, as articulações que se interligam, em sinergia, para produzir um movimento que permite ao ser humano caminhar. Por meio da análise biomecânica da marcha humana, é possível identificar algumas anomalias, que, ao serem analisadas, facilitam a criação de tecnologias que proporcionam um sistema de tratamento eficaz (ROSE; GAMBLE, 1998).

A marcha humana é considerada uma sequência de movimentos cíclicos, capaz de produzir um deslocamento em qualquer direção no eixo horizontal do ponto de vista humano. Nesses movimentos, enquanto um pé está apoiado sobre o solo, o outro está suspenso, realizando um movimento aéreo. A cada passo, invertem-se os membros

(ROSE; GAMBLE, 1998). Essa sequência de movimentos é considerada cíclica, pois o início do ciclo é o primeiro apoio de um dos pés sobre solo até o momento em que este mesmo pé realiza um novo contato. Assim, o ciclo da marcha humana é dividido em duas fases, apoio e balanço ou oscilação (MARCO SANZ, 2006).

Diversos estudos têm explorado a captura de movimento por meio de *hardwares* mais acessíveis, como o sensor Kinect (ALABBASI *et al.*, 2015; DAO *et al.*, 2015; MÜLLER *et al.*, 2017; SANG; YUAN; YU, 2021). O estudo de Alabbasi *et al.* (2015) investigou o problema de rastreamento esquelético de movimento do corpo humano com aplicação em reabilitação médica e treinamento esportivo usando o sensor Kinect, a fim de capturar e registrar os dados de rotação e posição entre as articulações um *personal trainer*. Esses dados foram combinados com uma exibição em tempo real de um personagem virtual em três dimensões (3D), mostrando seus movimentos na tela. Tendo como objetivo orientar um usuário a executar exercícios corretamente, o sistema comparava os movimentos do usuário final e os dados do profissional, exibindo uma pontuação na tela.

O estudo realizado por Dao *et al.* (2015) teve como objetivo a análise e visualização da marcha humana por meio dos dados capturados pelo sensor não intrusivo Kinect, classificando os dados com a máquina de vetores de suporte (*support vector machine* [SVM]) em anormal ou normal. Os autores capturaram os dados de indivíduos caminhando em direção ao sensor Kinect. Na sequência, processaram os dados com o filtro *B-spline* e os melhores ciclos da marcha capturada foram armazenados. Esses dados selecionados foram utilizados no treinamento da SVM, a fim de que esta se tornasse capaz de classificar as próximas entradas de dados capturados pelo sensor. Seus resultados obtiveram aproximadamente 88% de acurácia.

Na mesma linha de pesquisa, Müller *et al.* (2017) investigaram a validação de captura de movimento baseada em sensor Kinect aprimorada para avaliação de marcha. O objetivo desse trabalho era desenvolver um sistema de rastreamento do movimento baseado em sensores Kinect 2.0. Comparou-se a qualidade dos dados capturados pelo Kinect com os dados que foram capturados por um sistema de câmeras no ambiente de alta precisão Vicon. Utilizaram-se três Kinects para registrar a caminhada de indivíduos em sete metros. Os mesmos indivíduos repetiram a marcha com o sistema Vicon. Após a comparação dos dados, os autores propuseram uma configuração de câmera que contorna as desvantagens da captura de movimento pelo sensor Kinect, ocasionada pelos dados oclusos.

No trabalho de Sang, Yuan e Yu (2021), é apresentada uma estratégia que combina dados do Kinect com unidade de medida inercial com base no filtro de Kalman estendido para melhorar a precisão da estimativa de pose de robôs. Isso aprimora as soluções atuais, fortemente dependentes de dados visuais (mapa de profundidade), que podem ser afetados pelo ambiente e pelas condições de iluminação.

No contexto de robótica bioinspirada, realizou-se uma pesquisa relacionada à sintetização da marcha humana para projeto de um robô bípede. Com o auxílio do sensor de movimento Kinect 2.0 integrado ao *software* Simulink, da companhia MathWorks, foi feita a captura do movimento de 20 voluntários caminhando sobre uma esteira durante dois minutos (TAKEDA *et al.*, 2014). Os dados capturados foram armazenados em extensos arquivos numéricos.

A partir dos dados das marchas dos voluntários em arquivos, é possível aplicar técnicas de processamento de sinais e visualizá-los na forma de gráficos. Entretanto, a visualização dos dados nesse formato tem utilidade limitada para profissionais não experientes na análise de sinais. Sendo assim, percebeu-se a necessidade de visualização desses dados em animação de um modelo tridimensional de membros inferiores, de forma a possibilitar uma análise quanto à sua naturalidade/artificialidade por usuários leigos em análise de sinais. Lim *et al.* (2010) afirma que não existe uma definição adequada para o caminhar humano natural, entretanto, há um entendimento comum de como o ato de caminhar deve parecer para humanos, ou seja, ainda que não haja uma definição formal de como deve ser o caminhar humano, um indivíduo típico seria capaz de dizer se alguém não está caminhando de modo natural.

A marcha humana natural é conhecida por estar relacionada a parâmetros de marcha, como cadência, comprimento da passada e velocidade da caminhada. As variações de caminhada entre diferentes indivíduos ou de um mesmo indivíduo surgem de particularidades, como parâmetros de anatomia, mudanças na velocidade de caminhada e até alterações em calçados. Essas peculiaridades se sobrepõem ao padrão básico de locomoção bípedemas, geralmente, não afetam a percepção do indivíduo quanto à naturalidade ou não da marcha.

Nesse contexto, o presente trabalho busca realizar uma pesquisa de campo quanto à naturalidade/artificialidade da marcha humana a partir dos dados adquiridos anteriormente pelo Kinect. Para elaborar o instrumento de coleta de dados da pesquisa de campo, foram desenvolvidas animações a partir dos dados originais e também a partir de dados processados por diferentes filtros de sinais parametrizados. As animações de vista lateral e isométrica foram analisadas em uma pesquisa com a participação de 22 adultos, estudantes de graduação dos cursos de Engenharia de Computação e Sistemas de Informação. O instrumento de coleta de dados adotado foi o questionário *on-line*, com as animações tridimensionais incorporadas ao questionário na forma de vídeos. Os resultados da pesquisa de campo foram avaliados por meio da *mean opinion score* (MOS) e também foi realizada uma análise estatística dos resultados.

O objetivo deste trabalho é realizar a animação gráfica da marcha humana a partir um modelo tridimensional e de dados adquiridos pelo sensor Kinect, a fim de poder visualizar os dados capturados para uma posterior análise.

2 MATERIAIS E MÉTODOS

Os passos para a realização deste artigo foram os seguintes: criar um modelo esquelético com a utilização da interface de programação de aplicações (*application programming interface* [API]) OpenGL; selecionar os dados da marcha de um indivíduo adquiridos pelo Kinect no experimento anterior realizado em laboratório (TAKEDA *et al.*, 2014); animar o modelo tridimensional utilizando esses dados; suavizar os ruídos dos dados capturados com dois filtros; avaliar uma prévia dos movimentos da simulação gráfica; avaliar aspectos de qualidade dos dados adquiridos pelo sensor de movimentos Kinect por meio da aplicação de um questionário *on-line*; e analisar as respostas do questionário de acordo com a métrica MOS.

A linguagem de programação utilizada foi C++ e o ambiente de desenvolvimento utilizado foi o Visual Studio 2010. Para a aplicação gráfica, foi utilizada a API OpenGL e, para o desenvolvimento do modelo humano em três dimensões (3D) simplificado, utilizou-se a biblioteca Glut, que fornece funções para especificação e exibição de objetos primitivos, os quais, unidos e ajustados, permitem a criação de um modelo 3D simplificado de membros inferiores para as animações utilizando os dados de rotação das articulações adquiridos previamente pelo sensor Kinect. Para implementar as animações, foram utilizadas funções da biblioteca OpenGL (HEARN; BAKER; CARITHERS, 2004).

Para a implementação do modelo 3D simplificado, foram selecionados componentes da anatomia humana (ossos e juntas) correspondentes a um mapeamento entre os dados capturados pelo Kinect e as partes do modelo geométrico. Ainda que o modelo proposto seja uma simplificação dos membros inferiores humanos, pode realizar movimentos complexos, como o de agachar, no qual ocorre uma leve inclinação do joelho para frente e, como consequência, a tíbia e o fêmur também se movimentam, de forma que o ângulo interno entre o pé e a tíbia, bem como o ângulo entre a tíbia e o fêmur diminuam, como é mostrado na Figura 1.

No modelo da Figura 1(a), o ângulo α_1 é maior que o ângulo α_2 . Na Figura 1(b), ocorre o mesmo com os ângulos β_1 e β_2 , devido ao agachamento.

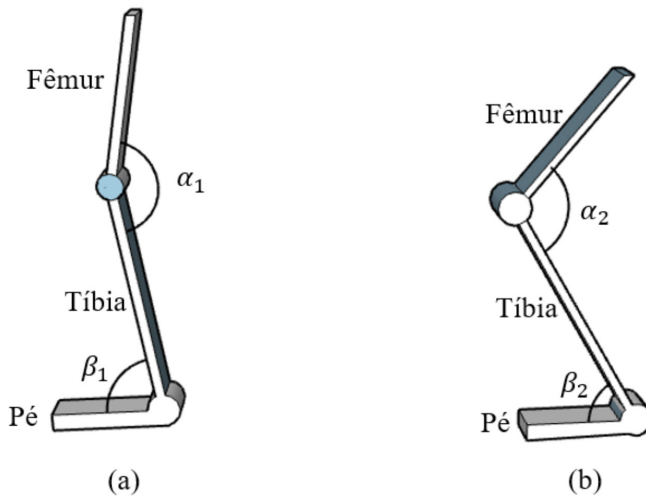


Figura 1 Modelo simplificado do membro inferior

Fonte: Elaborada pelos autores.

O modelo hierárquico tem uma posição inicial em estado de repouso e as transformações aplicadas às articulações desse objeto desencadeiam movimentos nas partes subordinadas hierarquicamente. A API OpenGL tem um mecanismo conhecido como pilhas de matrizes, adequado para criar objetos hierárquicos (SELLERS; WRIGHT JR.; HAEMEL, 2013). Esses movimentos são representados na forma de matrizes. No OpenGL, a função `glPushMatrix()` insere a matriz de transformação no topo da pilha, e a `glPopMatrix()` retira a matriz do topo da pilha, retornando a um estado anterior, sendo utilizadas para limitar as transformações de translação, rotação e escala a uma ou mais partes do modelo hierárquico.

A Figura 2 mostra o modelo anatômico simplificado e suas relações hierárquicas, correspondentes à parte inferior do corpo humano. A figura mostra, ainda, o posicionamento do sensor Kinect, simbolizado por um “olho”. Para exemplificar, se matriz relacionada ao fêmur esquerdo for alterada, ocorre uma alteração nas matrizes que estão subordinadas a esse nó (tíbia esquerda e o pé esquerdo).

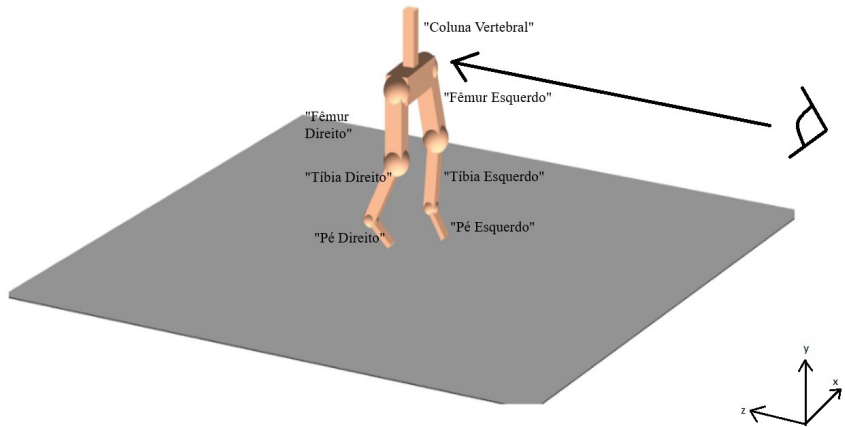


Figura 2 Modelo do membro inferior simplificado renderizado

Fonte: Elaborada pelos autores.

Na aquisição de dados realizada previamente, 20 participantes típicos com idade média de 22 anos, que relataram não apresentar qualquer comprometimento motor nos membros superiores e inferiores, caminharam em uma esteira instrumentalizada durante dois minutos. O recrutamento foi realizado a partir de um convite enviado por *e-mail* para os alunos regularmente matriculados ou que haviam cursado a disciplina de Processamento de Sinais, na instituição de ensino em que a pesquisa foi realizada. A aquisição dos dados foi realizada com o sensor *Kinect* posicionado em frente à esteira, no plano coronal. Os movimentos de rotações foram registrados na parte de trás das articulações. Os dados da captura de movimento dos voluntários com o Kinect (TAKEDA *et al.*, 2014) foram armazenados em uma matriz no Matlab. Essa variável representa a captura de seis ângulos dos membros inferiores de 20 voluntários gravados em 1 minuto e 53 segundos por voluntário, totalizando 3.593 *frames* por participante, a uma taxa de amostragem de 30 *frames* por segundo.

A Figura 3 apresenta o modelo com os seis ângulos capturados pelo sensor Kinect, denominados β_1 para representar a angulação entre o quadril esquerdo e o fêmur esquerdo; β_2 representando o ângulo entre o fêmur esquerdo a tibia esquerda; e β_3 , que representa o ângulo entre a tibia esquerda e o pé esquerdo, conforme a Figura 3(a). A Figura 3(b) mostra o modelo representado pelos ângulos β_4 entre o quadril direito e o fêmur direito; β_5 , que representa o ângulo entre o fêmur direito e a tibia direita; e β_6 , que representa a angulação entre a tibia direita e o pé direito.

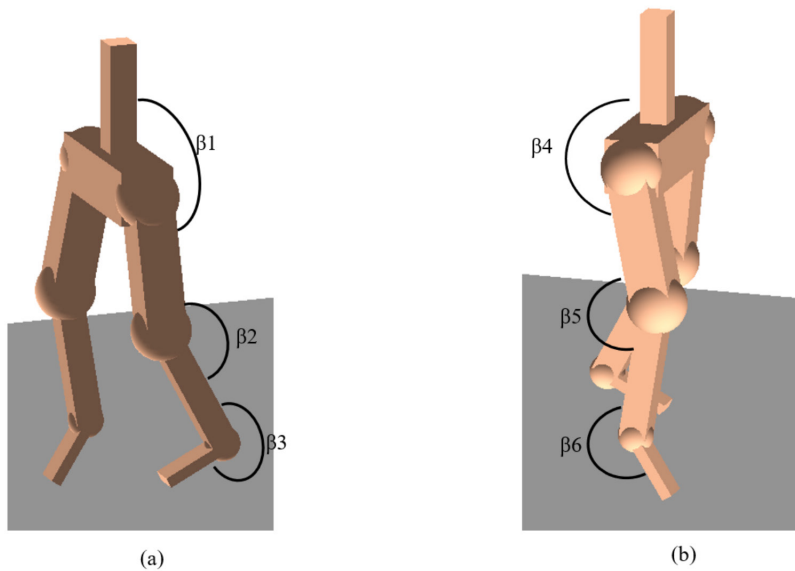


Figura 3 Ângulos capturados pelo Kinect

Fonte: Elaborada pelos autores.

A animação do modelo esquelético se baseou nos seis pontos capturados pelo Kinect, cada um deles tem um valor angular em rad/s. Para a animação do modelo 3D e processamento dos dados, somente os dados de um voluntário, escolhido de forma aleatória, foram utilizados. Foi implementada uma função em C++ para ler os dados e remover os 30 primeiros *frames*, pois a aquisição dos dados realizada pelo sensor Kinect utiliza aproximadamente os primeiros 30 *frames* para a calibração dos pontos no objeto a ser capturado.

Kinect é um sensor de jogos capaz de detectar e rastrear movimento, criado originalmente pela Microsoft para Xbox 360, Xbox One e computadores Windows. O Kinect utiliza seu sensor de profundidade para detectar e rastrear até 25 articulações do corpo humano e retorna suas localizações no espaço de coordenadas 3D a uma taxa de 30 *frames/s* (OSMAN; HEMBROFF; MATOS, 2018). A aquisição de dados pelo Kinect permite extrair o modelo de esqueleto de uma pessoa, independentemente das roupas e velocidade de caminhada. As trajetórias de ângulos articulares significativos, como quadril esquerdo, joelho esquerdo, quadril direito e joelho direito, são avaliadas a partir das coordenadas de esqueleto capturadas dos *frames* do vídeo de um indivíduo em movimento no campo de captura do Kinect (NANDY; CHAKRABORTY, 2015).

Os dados adquiridos pelo sensor Kinect, em decorrência da imprecisão do equipamento, podem ter falhas que levam a mudanças bruscas nos valores, chamados de ruídos no sinal. Segundo Osman, Hembroff e Matos (2018), diversos estudos mostraram que o Kinect apresenta imprecisão ao rastrear as articulações da parte inferior do corpo, especialmente em condições de rastreamento não ideais, como em situações com problemas de iluminação e oclusão das articulações durante a marcha. Dessa forma, torna-se necessário investigar se o processamento dos dados com a utilização de filtros de suavização para remover ou suavizar os ruídos pode melhorar a animação. Para isso, foram realizadas simulações com dois filtros de suavização, um no domínio do tempo (filtro da média móvel) e outro no domínio da frequência (filtro de Butterworth).

No processamento dos dados, foram realizadas iterações com o filtro de média móvel em diferentes animações, a partir do mesmo conjunto de dados brutos (BUTTERWORTH, 1930). Para comprovar essa hipótese, o método do estudo foi reunir um vídeo sem filtros (dados brutos) e os vídeos com filtros no questionário *on-line* e comparar as respostas dos participantes quanto à sua percepção de naturalidade/artificialidade das animações.

Os dados de filtro de média móvel tiveram a variação de dois parâmetros, a saber: o número de repetições em 1, 3 e 5 e a janela do filtro de 10, 15, 20 e 25 referentes aos parâmetros usados no filtro de média móvel. Desse modo, foram geradas 12 combinações de parâmetros (número de repetições, tamanho da janela) além dos dados originais sem processamento. Com a utilização do filtro de média móvel, foram geradas 24 animações, sendo 12 da perspectiva isométrica e 12 da lateral, para posteriormente serem inseridas na pesquisa de campo. No processamento dos dados com o filtro de Butterworth, foi feita a variação de dois parâmetros, a frequência de corte de 0,5, 1,0 e 2,0 e a ordem do filtro (3 ou 5). Foram gerados 12 vídeos, sendo seis na perspectiva isométrica e seis na vista lateral. Além das animações correspondentes ao uso dos filtros parametrizados, foram incorporadas à pesquisa de campo animações a partir das perspectivas isométrica e lateral contendo os dados originais sem processamento.

Na Figura 4, nota-se que os dados originais têm ruídos provenientes da captura do movimento. Os dados utilizados nessa imagem são do quadril esquerdo.

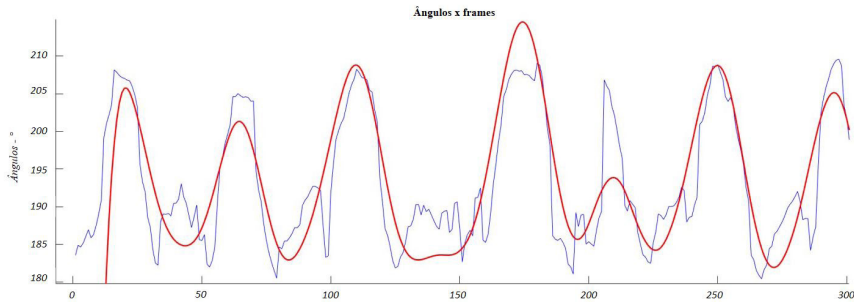


Figura 4 Gráfico dos dados com processamento e sem processamento

Fonte: Elaborada pelos autores.

Nota: o azul representa os dados originais e o vermelho, os dados processados com tamanho de janela igual a 15 e número de repetições igual a três.

Antes de iniciar a pesquisa de campo, analisamos as 40 animações resultantes do processamento com os filtros, em busca de artefatos que evidenciassem um ciclo da marcha humana mais natural, concluindo que a animação realizada com o filtro de média móvel com janela igual a 15 e número de repetições igual a três é o mais natural. Todas outras animações geradas com parâmetros diferentes, tanto as geradas com filtro de média móvel e filtro de Butterworth quanto as animações com os dados brutos (sem filtro), foram avaliadas pelos participantes da pesquisa de campo. Os participantes da aquisição de dados pelo Kinect foram excluídos da pesquisa de campo. O recrutamento para essa etapa da pesquisa foi realizado por meio de envio de um *e-mail* para a comunidade acadêmica do *campus* da instituição de ensino em que a pesquisa foi realizada.

A partir das animações correspondentes às diversas combinações de parâmetros dos dois filtros selecionados e as animações originais, sem processamento de sinal, foram gerados 40 vídeos das animações nomeados de #1 a #40, apresentados em ordem aleatória a um público externo, a fim de avaliar as animações quanto à naturalidade/artificialidade do movimento.

O questionário *on-line* utilizado na pesquisa de campo foi elaborado da seguinte forma: primeiramente, introduziu-se brevemente a pesquisa contendo o objetivo, número de vídeos para análise e tempo esperado de participação. Também foram incluídas algumas perguntas para levantar o perfil dos respondentes. Não foram coletados dados sensíveis ou que pudessem identificar os voluntários da pesquisa. As questões buscaram captar a percepção dos respondentes quanto à naturalidade/artificialidade de cada uma das animações. A primeira pergunta foi elaborada com a MOS. Seguindo essa técnica, realizou-se uma adaptação da escala para avaliar a

qualidade da marcha humana nessa pesquisa. A escala proposta é apresentada na Tabela 1.

TABELA 1
Escala para avaliação da marcha humana

Avaliação	Etiqueta
5	Completamente natural sem movimentos artificiais
4	Natural com movimentos artificiais esporádicos
3	Natural com movimentos artificiais frequentes
2	Artificiais com movimentos naturais esporádicos
1	Completamente artificial e inaceitável

Fonte: Elaborada pelos autores.

A segunda pergunta, com escala de 0 a 10, avalia de forma simples e subjetiva a qualidade da marcha humana no vídeo. Uma terceira pergunta foi inserida de acordo com a resposta da primeira. Se fosse escolhida a opção 1 – completamente artificial e inaceitável –, uma nova pergunta dissertativa era apresentada ao entrevistado, a fim de coletar mais dados a respeito de sua percepção.

Os resultados obtidos pela pesquisa de campo foram submetidos a uma análise utilizando a métrica MOS. Essa técnica é utilizada na área da engenharia de telecomunicações com a finalidade de avaliar a qualidade da voz subjetivamente e pode ser aplicada à pesquisa em outras áreas que envolvam sinais que dependem da opinião das pessoas (JUANG; SONDHI; RABINER, 2003). Para analisar os resultados, é calculada a média aritmética dos valores de cada vídeo/animação em uma escala utilizando Equação 1.

$$MOS = \frac{\sum_{N=1}^n R_n}{N} \quad (1)$$

em que:

R = avaliação individual para os determinados grupos;

N = número de grupos.

Os resultados de MOS foram comparados por meio da análise de variância *one way*, que indica a diferença entre os vídeos em função da variação de cada um.

3 RESULTADOS E DISCUSSÃO

A pesquisa de campo foi realizada em um período de 15 dias e os dados foram compilados e analisados. Ao todo, participaram da pesquisa de campo sobre naturalidade/artificialidade das animações 22 indivíduos, sendo: 76,2% do público composto por homens e 23,8%, por mulheres, todos estudantes de graduação dos cursos de Engenharia de Computação e Sistemas de Informação.

Na questão 1, relacionada à escala de naturalidade/artificialidade, todas as 40 animações foram exibidas em ordem aleatória a partir de duas perspectivas diferentes, uma vista isométrica e outra lateral.

Os respondentes avaliaram os 40 vídeos resultantes quanto à naturalidade/artificialidade em uma escala de 0 a 5. Na pesquisa de campo, os vídeos #33 e #39 foram os que obtiveram maior nota MOS de naturalidade. Esses vídeos também foram os melhores avaliados de acordo com a análise preliminar realizada pelos autores do estudo. Quanto ao vídeo #33, cerca de 57,1% dos entrevistados relataram que os movimentos são completamente naturais sem movimentos artificiais, já no vídeo #39, em torno de 47,6% relataram que os movimentos são naturais com movimentos artificiais esporádicos. Os vídeos #33 e #39 correspondem às perspectivas isométrica e lateral, respectivamente, da marcha percebida como a mais natural pelos participantes do estudo.

Essa divergência de percepção pode ter sido ocasionada por se tratar de uma visão lateral, o que pode ter evidenciado alguns detalhes não observados na imagem de perspectiva isométrica.

Os vídeos #11 e #18 obtiveram as piores MOS. Quanto ao vídeo #11, cerca de 71,4% avaliaram a marcha sintética como completamente artificial e inaceitável, enquanto, para o vídeo #18, em torno de 66% avaliaram a animação como completamente artificial. Ambos os vídeos foram gerados a partir da animação dos dados originais capturados pelo sensor Kinect. Esse resultado já era esperado, uma vez que o sensor apresenta limitações devidas à oclusão no momento da captura dos dados (MÜLLER *et al.*, 2017). Essas limitações de captura do sensor Kinect relatadas na literatura e observadas pelos pesquisadores na análise preliminar dos sinais adquiridos foram decisivas para evidenciar a necessidade de melhoria dos sinais com o processamento de filtros para remoção dos ruídos. Na segunda questão, os respondentes avaliaram a qualidade geral do vídeo, utilizando uma escala numérica de 0 a 10. Novamente, os vídeos #33 e #39 obtiveram as maiores notas absolutas segundo a análise de MOS.

Na questão 3, relacionada à artificialidade, os 22 respondentes classificaram pelo menos um dos vídeos como artificial. Desses, nove relataram o que consideraram artificial no vídeo. As respostas obtidas foram: “Porque os movimentos não correspondem ao de uma caminhada”, “Movimentos não naturais descompassados”, “Pois está em desacordo com o andar de um ser humano”, “Pois a perna não flexiona para trás”, “Movimento não natural e desnecessário dos pés”, “Muita vibração”, “Pois o pé está tremendo muito”, “Movimento dos pés muito travado”, “Padrão de movimento fora do normal” e “Os pés estão mexendo exageradamente”. Percebe-se pelos relatos que os pés são os fatores que mais influenciam na avaliação da marcha. Portanto, os relatos corroboram os estudos realizados sobre a marcha humana.

Os dados obtidos na pesquisa de campo foram submetidos a uma análise de variância com nível de significância de 0.05. O resultado calculado de $p\text{-value}$ é $9,99201 \times 10^{-16} < 0,05$ ($p\text{-value} < \alpha$), portanto a hipótese nula é rejeitada, ou seja, as diferenças entre algumas das médias de alguns grupos são suficientes para ser estatisticamente significativas.

Os dois melhores vídeos pela análise de MOS foram #33 e #39, que se referem ao vídeo sintetizado com o filtro de média móvel, com janela igual a 15 e número de repetições igual a três, nas perspectivas isométrica e lateral, respectivamente. Utilizando o teste de Tukey-Kramer, foi possível comparar a diferença entre os vídeos #33 e #39 em relação aos demais vídeos. A Tabela 2 apresenta a médias dos pares que são significativamente diferentes para o vídeo #33.

Os vídeos #3, #8, #11, #12 e #13 são vídeos das animações processadas com o filtro de média móvel na perspectiva isométrica. Esses vídeos obtiveram as menores métricas MOS. O vídeo #18 contém a animação original, com dados ruidosos, gerando uma animação de qualidade inferior. Os vídeos #14, #18, #19, #21, #24, #27 e #31 foram gravados a partir da vista lateral, sendo incluídos nos pares de vídeos com médias significativas, pois contêm valores baixos de MOS e uma grande diferença na análise de variância. Os vídeos #27 e #31 contêm animações provenientes do processamento do filtro de Butterworth.

TABELA 2

Diferença entre as médias significativas do vídeo #33

Vídeo			Vídeo			Diferença
Animação	Vista	Filtro	Animação	Vista	Filtro	
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#3	Isométrica	Média móvel (janela = 20 e repetições = 5)	1,50
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#8	Isométrica	Média móvel (janela = 10 e repetições = 1)	1,45
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#11	Isométrica	Original	2,68
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#12	Isométrica	Média móvel (janela = 25 e repetições = 5)	1,82
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#13	Isométrica	Média móvel (janela = 25 e repetições = 3)	1,36
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#14	Lateral	Média móvel (janela = 10 e repetições = 1)	2,05
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#18	Lateral	Original	2,77
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#19	Lateral	Média móvel (janela = 15 e repetições = 1)	1,55
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#21	Lateral	Média móvel (janela = 25 e repetições = 5)	1,64
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#24	Lateral	Média móvel (janela = 20 e repetições = 5)	1,36
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#27	Lateral	Filtro de Butterworth (ordem = 5 e frequência de corte = 2.0)	1,59
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#31	Lateral	Filtro de Butterworth (ordem = 3 e frequência de corte = 2.0)	1,55
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#35	Isométrica	Filtro de Butterworth (ordem = 5 e frequência de corte = 2.0)	1,55
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#36	Isométrica	Filtro de Butterworth (ordem = 3 e frequência de corte = 2.0)	1,27
#33	Isométrica	Média móvel (janela = 15 e repetições = 3)	#37	Isométrica	Filtro de Butterworth (ordem = 5 e frequência de corte = 0.5)	1,59

Fonte: Elaborada pelos autores.

TABELA3

Diferença entre as médias significativas do vídeo #39

Animação	Vídeo		Animação	Vídeo		Diferença
	Vista	Filtro		Vista	Filtro	
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#3	Isométrica	Média móvel (janela = 20 e repetições = 5)	1,45
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#8	Isométrica	Média móvel (janela = 10 e repetições = 1)	1,41
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#11	Isométrica	Original	2,64
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#12	Isométrica	Média móvel (janela = 25 e repetições = 5)	1,77
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#13	Isométrica	Média móvel (janela = 25 e repetições = 3)	1,32
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#14	Lateral	Média móvel (janela = 10 e repetições = 1)	2,00
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#18	Lateral	Original	2,73
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#19	Lateral	Média móvel (janela = 15 e repetições = 1)	1,50
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#21	Lateral	Média móvel (janela = 25 e repetições = 5)	1,59
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#24	Lateral	Média móvel (janela = 20 e repetições = 5)	1,32
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#27	Lateral	Filtro de Butterworth (ordem =5 e frequência de corte = 2.0)	1,55
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#31	Lateral	Filtro de Butterworth (ordem = 3 e frequência de corte = 2.0)	1,50
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#35	Isométrica	Filtro de Butterworth (ordem = 5 e frequência de corte = 2.0)	1,50
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#36	Isométrica	Filtro de Butterworth (ordem = 3 e frequência de corte = 2.0)	1,23
#39	Lateral	Média móvel (janela = 15 e repetições = 3)	#37	Isométrica	Filtro de Butterworth (ordem = 5 e frequência de corte = 0.5)	1,55

Fonte: Elaborada pelos autores.

A Tabela 3 contém informações similares às da Tabela 2, por se tratar de um vídeo (#39) que foi gravado com o filtro de média móvel com janela de 15 e número de repetições igual a três na perspectiva vista lateral. Nesse ângulo de visualização, podem ser observados artefatos não visíveis no ângulo de vista isométrica – caso do vídeo #33.

4 CONSIDERAÇÕES FINAIS

Este trabalho apresentou os resultados de uma pesquisa de campo quanto à naturalidade/ artificialidade da marcha humana, a partir de dados capturados por sensor. Os dados foram processados com a utilização de dois filtros de suavização, filtro de média móvel e filtro de Butterworth, e animados em um modelo 3D simplificado do esqueleto dos membros inferiores humanos.

O sensor utilizado, Kinect, tem um sistema de captura ótico e acústico, sendo um equipamento de baixo custo. Entretanto, o Kinect apresenta limitações durante a aquisição dos dados, muitas vezes ocasionadas pela obstrução de alguma parte do objeto durante o movimento. Diante disso, o sensor tem sido pouco explorado para a análise da marcha humana.

Este estudo mostra que é possível utilizar um sistema de baixo custo para a captura de movimento e processar os dados com técnicas de processamento de sinais, removendo os ruídos provenientes de falhas e imprecisões durante a aquisição dos dados pelo sensor. Como resultado dessa avaliação da marcha humana, conclui-se que os dados capturados podem ser processados com o auxílio de filtros que eliminam os ruídos provenientes da captura de movimento pelo sensor e esses dados pré-processados podem ser avaliados conforme a métrica MOS. Participantes de uma pesquisa de campo avaliaram as animações quanto à artificialidade/naturalidade da marcha, como resultado, obteve-se uma avaliação MOS máxima de 4,48. Utilizando a análise de variância, os vídeos #33 e #39 foram os mais naturais, contendo médias significativas em relação a outros vídeos.

Os resultados preliminares deste estudo sugerem que o processamento dos dados adquiridos por sensores de baixo custo, como o Kinect, pode auxiliar na percepção da marcha sintética, sendo adequado para estudos que envolvem a análise da marcha humana, como a construção de modelos robóticos que precisem de dados para se locomover, ou ainda uma análise biomédica de um indivíduo em fase de recuperação pós-trauma, de modo a auxiliar profissionais de saúde no tratamento do paciente.

GRAPHIC ANIMATION OF HUMAN GAIT WITH DATA COLLECTED FROM THE KINECT SENSOR

Abstract

The analysis of human gait from biometric data has applications in areas of safety, bio-inspired robotics, and health. Low-cost motion sensors – for instance, Kinect – have enabled the acquisition of biometric gait data in terrestrial environments. However, these devices have limitations that can impact the quality of the acquired data. In this scenario, different signal processing techniques can be applied to reduce noise. Visualization of this data, whether original or processed, is often performed in the form of graphs, having limited utility for professionals not experienced in signal analysis. In this sense, the visualization of gait data in a three-dimensional model can contribute to better decision-making by professionals, especially in healthcare. This work aims to perform the graphic animation of the human gait in a three-dimensional model from data collected by the Kinect 2.0 sensor. To reduce the noise of the acquired data, pre-processing was performed with the application of moving average and Butterworth filters. Videos of the animations from isometric and side views were created and included in the online questionnaire to evaluate the artificiality/naturalness of the animation, using the mean opinion score (MOS) technique. A total of 22 participants, undergraduate students of Computer Engineering and Information Systems, answered the online questionnaire. The one-way analysis of variance (ANOVA) statistical test showed that videos from the isometric and side perspectives processed with a moving average filter (window = 15 and repetitions = 3) that obtained higher MOS metric values were evaluated as significantly more natural than the others, processed or not.

Keywords: Animation. Kinect. Human march.

REFERÊNCIAS

ALABBASI, H. *et al.* Human motion tracking & evaluation using Kinect V2 sensor. *In: E-HEALTH AND BIOENGINEERING CONFERENCE*, 5., 2015. *Proceedings* [...]. New York: IEEE, 2015. DOI 10.1109/EHB.2015.7391465

BORENSTEIN, G. *Making things see: 3D vision with Kinect, processing, Arduino, and MakerBot*. [s. l.]: O'Reilly Media, 2012.

BUTTERWORTH, S. On the theory of filter amplifiers. *Experimental Wireless and the Wireless Engineer*, v. 7, n. 6, p. 536-541, 1930.

DAO, N.-L. *et al.* Kinect-based non-intrusive human gait analysis and visualization. *In: E-HEALTH AND BIOENGINEERING CONFERENCE*, 5., 2015. *Proceedings* [...]. New York: IEEE, 2015. DOI 10.1109/MMSP.2015.7340804

HEARN, D.; BAKER, M.; CARITHERS, W. *Computer graphics with Open GL*. Upper Saddle River, NJ: Pearson Prentice Hall, 2004. v. 3.

JUANG, B. H.; SONDHAI, M.; RABINER, L. R. Digital speech processing. *In: MEYERS, R. A. (Ed.). Encyclopedia of physical science and technology*. 3. ed. New York: Academic Press, 2003. p. 485-500. ISBN 978-0-12-227410-7. Disponível em: <https://www.sciencedirect.com/science/article/pii/B0122274105001782>. Acesso em: 10 mar. 2022.

LIM, H. B. *et al.* Natural gait parameters prediction for gait rehabilitation via artificial neural network. *In: INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS*, 2010. *Proceedings* [...]. New York: IEEE, 2010. DOI 10.1109/IROS.2010.5650311

MARCO SANZ, C. *Cinesiologia de la marcha humana normal*. 2006. Disponível em: <http://wzar.unizar.es/acad/cinesio/Documentos/Marcha%20humana.pdf>. Acesso em: 10 mar. 2022.

MIYAWAKI, K.; SATO, A.; KIZAWA, S. Evaluating motion capture accuracy for gait analysis. *In: INTERNATIONAL SYPOSIUM ON MICRO-NANOMECHATRONICS AND HUMAN SCIENCE. Proceedings* [...], p. 1-4, 2017. DOI 10.1109/MHS.2017.8305163

MÜLLER, B. *et al.* Validation of enhanced Kinect sensor based motion capturing for gait assessment. *PloS One*, San Francisco, v. 12, n. 4, e0175813, 2017.

NANDY, A.; CHAKRABORTY, P. A new paradigm of human gait analysis with Kinect. *In: EIGHTH INTERNATIONAL CONFERENCE ON CONTEMPORARY COMPUTING*, 2015. *Proceedings* [...]. New York: IEEE, 2015. DOI 10.1109/IC3.2015.7346722

OSMAN, S. E. K.; HEMBROFF, G. C.; MATOS, M. A. A novel Kinect-based algorithm for inferring the position of the lower body joints using human gait pattern. *In: INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS*, 11., 2018. *Proceedings* [...]. New York: IEEE, 2018. DOI 10.1109/ICHI.2018.00042

ROSE, J.; GAMBLE, J. G. *Marcha humana*. São Paulo: Premier, 1998.

SANG, X.; YUAN, Z.; YU, X. Inertial-Kinect fusion for robot navigation based on the extended Kalman filter. *In: ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY*, 47., 2021. *Proceedings* [...]. New York: IEEE, 2021.

SELLERS, G.; WRIGHT JR., R. S.; HAEMEL, N. *OpenGL Superbible: comprehensive tutorial and reference*. Boston, MA: Addison-Wesley, 2013.

SILVA, R. G. da *et al.* Application of a demand forecasting model in a rental company of billiard tables. *Journal of Engineering and Technology for Industrial Applications*, v. 4, n. 15, p. 53-58, 2018. DOI 10.5935/2447-0228.20180047

TAKEDA, R. Y. *et al.* Identificação biométrica de indivíduos baseada em parâmetros dinâmicos extraídos da marcha humana. *In*: SIMPÓSIO DE PROCESSAMENTO DE SINAIS DA UNICAMP, 5., 2014. *Anais [...]*. Campinas: Unicamp, 2014.

Contato

Edmo de Oliveira Leite
edmo.leite@aluno.ufop.edu.br

Tramitação

Recebido em outubro de 2021.
Aprovado em dezembro de 2021.