

# Tangible Beats: uma Interface Tangível para Expressão e Exploração Musical

Alexandre M. F. de Sousa e Carlos H. Morimoto  
 Departamento de Ciência da Computação  
 Universidade de São Paulo, Brasil  
 {alemart, hitoshi}@ime.usp.br

## RESUMO

Este artigo descreve o desenvolvimento do Tangible Beats, uma interface tangível para criação e exploração de padrões musicais. Com ela, os usuários elaboram, colaborativamente, padrões musicais manipulando objetos coloridos dispostos numa superfície plana em forma de grade. Músicos amadores podem expressar sua musicalidade e criatividade. Nosso protótipo usa uma câmera Kinect para detectar os artefatos coloridos e um projetor de vídeo para fornecer retorno visual.

## Palavras-chave

Interface tangível; expressão e exploração musical colaborativa; Kinect.

## ABSTRACT

This paper describes the development of Tangible Beats, a tangible musical interface for the creation and collaborative exploration of musical patterns. Tangible Beats allows users to collaboratively play musical patterns by manipulating colored tokens on a flat grid surface. Inexperienced players can explore and express their musicality. Our prototype uses a Kinect camera to detect the position of the colored objects. Visual feedback is provided using a video projector.

## Author Keywords

Tangible interface; collaborative music expression and exploration; Kinect.

## ACM Classification Keywords

H.5.2. User Interfaces: Input devices and strategies; Interaction styles.

## General Terms

Design, Experimentation.

## INTRODUÇÃO

A criação e exploração musical utilizando instrumentos convencionais geralmente exige certo treinamento, requerindo a realização de movimentos físicos específicos. Com a evolução dos sistemas computacionais para performances

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. IHC'14, Brazilian Symposium on Human Factors in Computing Systems. October 27-31, 2014, Foz do Iguaçu, PR, Brazil. Copyright 2014 SBC. ISSN 2316-5138 (pendrive). ISBN 978-85-7669-291-1 (online).



Figura 1: Uma instalação típica do Tangible Beats.

musicais, pessoas com menor habilidade, treinamento e conhecimento musical passaram a ser capazes de participar do processo criativo e exploratório. O grande interesse e a rápida evolução da área são evidenciadas pela criação, a partir de um workshop realizado no ACM CHI em 2001, da conferência internacional “Novas Interfaces Para Expressão Musical” (NIME - New Interfaces for Musical Expression).

Dentre as inúmeras alternativas sugeridas, o uso de interfaces tangíveis [2] oferece simplicidade de uso e facilidade de aprendizagem quando comparado a interfaces gráficas tradicionais, além de permitir e promover a colaboração entre vários usuários de forma natural. Para dar uma ideia da evolução dessas interfaces e sua relação com este trabalho, vamos descrever brevemente as interfaces do Block Jam [8], do Reactable [3] e do Tactus [6].

O Block Jam, desenvolvido por Newton-Dunn, *et al.* [8], dispõe de uma interface tangível composta por 26 artefatos físicos (blocos), cada um com uma função musical distinta. Um sequenciador coleta as informações funcionais e topológicas dos blocos e as interpreta na forma de estruturas e frases musicais. Cada bloco do Block Jam funciona como um complexo dispositivo eletrônico capaz de se conectar magneticamente com outros blocos, comunicar-se com os demais blocos por meio de uma via de dados, processar informações e exibir o seu estado numa matriz de LEDs.

Talvez a interface musical tangível mais conhecida pela sua popularidade seja o Reactable, desenvolvida por Jordà *et al.* [3]. O Reactable utiliza uma mesa interativa circular de 85 cm de diâmetro e 90 cm de altura. A superfície da mesa é

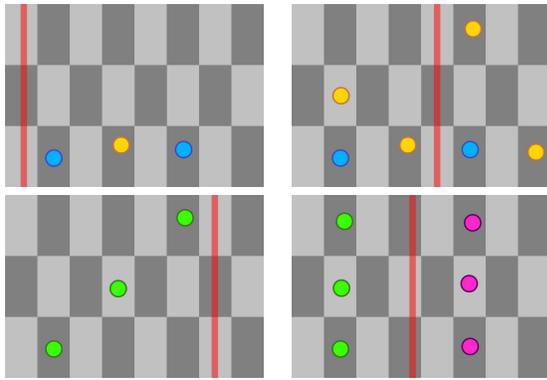


Figura 2: Exemplos de composições no Tangible Beats.

retro projetada para exibir informações gráficas aos usuários. A forma circular foi escolhida para facilitar a colaboração entre os participantes. Uma grande vantagem do Reactable sobre o Block Jam é a simplificação dos blocos. No caso do Reactable, marcas fiduciais são coladas nas faces dos blocos, e um sistema de visão computacional detecta e interpreta as marcas. Cada marca corresponde a uma funcionalidade distinta. O usuário conecta e desconecta os blocos aproximando-os e distanciando-os, respectivamente. O Reactable atingiu grande popularidade, sendo utilizado em shows internacionais pela cantora islandesa Björk. Além disso, os inventores do Reactable fundaram uma companhia para a comercialização do produto <sup>1</sup>.

A ideia de conectar objetos também foi utilizada também por Mann *et al.* [6] no desenvolvimento do sistema Tactus. O Tactus permite que qualquer estrutura retangular em forma de grade, onde algumas peças possam ficar presas, seja transformada num sequenciador musical de passo. Um sistema de visão computacional identifica as posições livres e ocupadas da grade. A disposição das peças numa mesma coluna define uma nota, e as colunas são “tocadas” em sequência, repetidamente, num processo de varredura. A música tocada é alterada quando novas peças são introduzidas ou removidas da grade.

Neste artigo descrevemos o desenvolvimento do Tangible Beats. Uma câmera Kinect é utilizada para a detecção de objetos em contato com uma superfície plana. A câmera é acoplada a um projetor de vídeo, que fornece retorno visual. Este sistema tangível é demonstrado na forma de um sequenciador musical de passo como o Tactus, mas que oferece uma experiência diferente devido à presença de *feedback* visual animado e de uma área interativa maior, facilitando a colaboração.

### INTERAÇÃO TANGÍVEL

A figura 2 ilustra quatro composições feitas no Tangible Beats. O fundo da cena apresenta uma grade retangular com 3 linhas e 8 colunas. A linha vermelha vertical é chamada de linha de varredura. O usuário pode posicionar objetos coloridos em qualquer célula livre da grade. Discos coloridos são

<sup>1</sup><http://www.reactable.com/>. Último acesso: 27 de julho de 2014.

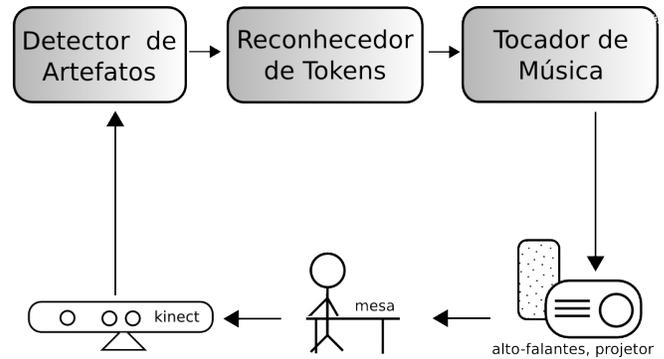


Figura 3: Visão geral do sistema.

projetados ao redor dos objetos como forma de informar ao usuário que o sistema os detectou. Quando a linha de varredura ultrapassa um objeto, o som correspondente é tocado. Observe que vários sons podem ser tocados simultaneamente. Peças de cores diferentes representam instrumentos distintos, e a altura do som gerado varia com a altura do objeto na coluna, sendo grave na linha inferior e agudo na linha superior.

### IMPLEMENTAÇÃO

A implementação do protótipo, representada na figura 3, traz consigo três subsistemas: Detector de Artefatos, Reconhecedor de Tokens e Tocador de Música. O Detector de Artefatos transforma uma superfície plana não-instrumentalizada numa região sensível a toques. Tal funcionalidade provê meios de determinar a posição das peças colocadas sobre a mesa. O Reconhecedor de Tokens, por outro lado, analisa as imagens obtidas pela câmera para determinar a cor das peças. Uma vez determinadas a posição e a cor de cada token, entra em cena o Tocador de Música, que provê retorno visual e sonoro ao usuário. Do ponto de vista de hardware, a instalação do protótipo requer: um laptop, uma câmera RGB-D (em nosso protótipo utilizamos um Microsoft Kinect), alto-falantes e um pequeno projetor de vídeo de 400 lm preso a um tripé.

#### Detector de Artefatos

O módulo de detecção de artefatos é dividido em dois componentes: Detector de Toques e Rastreador dos Artefatos.

Usando um Kinect, Wilson propôs uma técnica de processamento de imagens que detecta toques numa mesa não-instrumentalizada [9]. Apesar de os resultados não serem tão acurados se comparados às tecnologias multi-toque capacitivas, a abordagem por ele apresentada se mostra apropriada para aplicações diversas - com o benefício adicional de se poder explorar a forma e a textura dos objetos que tocam a superfície. Xiao *et al.* [10] utilizaram uma técnica similar no WorldKit, um sistema de realidade aumentada aplicado a superfícies do dia-a-dia (ex: móveis de uma sala de estar ganham componentes virtuais com as quais se poder interagir).

A técnica utilizada neste sistema é baseada no trabalho de ambos, com a diferença de que aplicamos uma máscara na imagem de profundidade fornecida pela câmera para restringir a detecção de toques a uma superfície plana (como uma mesa).

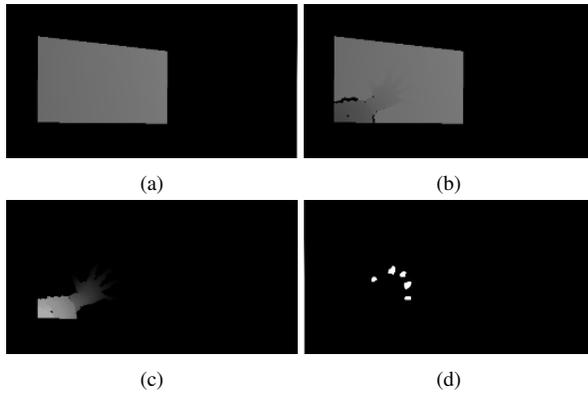


Figura 4: (a) Modelo de fundo. (b) Uma imagem de profundidade. (c) Imagem após subtração de fundo. (d) Blobs resultantes.

A figura 4 mostra as operações aplicadas na imagem de profundidade. O centro de massa das componentes resultantes indica a posição  $(x, y)$  de um ponto de toque no sistema de coordenadas do sensor de profundidade. Tais coordenadas são transformadas para um espaço normalizado  $[0, 1] \times [0, 1]$  por meio de uma matriz de homografia, computada utilizando um procedimento padrão de calibração de quatro pontos [1].

#### Rastreador dos Artefatos

Detectar os pontos de toque na superfície plana equivale a determinar a posição dos tokens sobre a mesa. Isto feito, o sistema rastreia cada um dos tokens no tempo. O algoritmo de *tracking* computa, de maneira gulosa, uma associação entre as posições dos artefatos encontrados na imagem atual fornecida pelo Kinect e as posições dos artefatos detectados no quadro anterior. Como sugerido em Kramer, *et al.* [4], novos IDs são atribuídos a todos os artefatos da imagem atual. Então, prediz-se a nova posição de cada artefato do quadro anterior utilizando posição e velocidades anteriores. Sob a suposição de que a nova localização dos tokens é encontrada próxima à posição predita, efetuamos a associação repassando o ID do token do quadro anterior ao token correspondente do novo quadro. O algoritmo termina com os todos os novos artefatos recebendo novos IDs.

#### Reconhecimento de Tokens

Após detectar a posição dos tokens dispostos sobre a mesa, o sistema precisa reconhecer a cor dos objetos. Tal operação se dá em duas etapas: extração e classificação de cor. A câmera RGB incluída no Kinect é utilizada para extrair dados de cor. As imagens RGB e de profundidade do Kinect são inicialmente alinhadas.

#### Extração de cor

Dada a posição  $s' = [s'_x \ s'_y]^t$  de um token na imagem de profundidade, a abordagem descrita por Kramer, *et al.* [4] é utilizada para computar sua posição correspondente  $s = [s_x \ s_y \ s_z]^t$  no espaço físico tridimensional. Para um raio  $r$  pré-definido, considera-se a seguinte esfera:

$$\gamma : (x - s_x)^2 + (y - s_y)^2 + (z - s_z)^2 = r^2$$

Ao tomar as cores de uma amostra de pontos em  $\gamma$ , computa-se suas respectivas cromacidades no espaço CIELAB, o que resulta num conjunto  $C$  de pares de números que representa a cor do token.

#### Classificação de cor

Inicialmente, cada par cromático  $p_i \in C$  é analisado individualmente. A partir da observação de que o cluster cromático  $C$  tem formato similar a uma elipse, decidimos ampliar o *elliptical boundary model* de Lee e Yoo, originalmente proposto como um classificador binário de cor de pele [5].

Para cada possível categoria  $j$  de cor, computa-se o conjunto:

$$M_j = \{ p_i \in C \mid (p_i - \phi_j)^t \Lambda_j^{-1} (p_i - \phi_j) < \theta \}$$

$\theta$  é um valor especificado empiricamente e  $\phi_j \in \mathbb{R}^2$  e  $\Lambda_j \in \mathbb{R}^{2 \times 2}$  são os modelos do parâmetro computados de acordo com [5]. Definindo o volume sinalizado  $v_j$  como sendo o determinante  $|\Lambda_j|$ , o cluster  $C$  é classificado como a cor  $j$  que maximiza a densidade:

$$d_j = \frac{|M_j|}{|v_j|}$$

Se  $|M_j| = 0$  para todo  $j$ , então  $C$  é classificado como cor de pele. Adicionalmente, se um mesmo token for classificado como  $j_k$  por  $m$  ou mais quadros consecutivos, então ele será sempre classificado como  $j_k$ .

#### Tocador de Música

Os subsistemas descritos anteriormente abordam a componente de visão computacional deste trabalho. Eles fornecem um conjunto de triplas  $(x_i, y_i, c_i)$  que indicam a posição e a cor dos tokens. Munido de tais dados, o Tocador de Música produz estímulos sonoros e visuais para o usuário. A grade é renderizada com cores neutras, de modo a preservar a matiz dos tokens coloridos mesmo na presença de um projetor de vídeo.

O Tocador de Música foi desenvolvido utilizando openFrameworks, um toolkit C++ multi-plataforma que provê diversos recursos para programadores de arte interativa [7]. Como a linha de escaneamento varre a mesa no eixo-x, faz-se necessário discretizar as posições dos tokens, dado que elas residem num espaço contínuo  $[0, 1] \times [0, 1]$ . Foi utilizada a seguinte função de discretização:

$$D_w(x) = \frac{\lfloor wx \rfloor}{w} + \frac{1}{2w}$$

onde  $w \in \mathbb{N}^*$  é o número de células da grade no eixo horizontal. Sempre que a linha de escaneamento, caracterizada por uma variável  $s \in [0, 1]$ , ultrapassa um token, o som correspondente é tocado.

## RESULTADOS

O protótipo do Tangible Beats, mostrado na figura 1, foi apresentado numa exibição de arte em nossa universidade. Durante a exibição, pessoas de todas as idades puderam experimentar o sistema livremente (ver figura 5). Durante o evento, notamos que as pessoas “descobriam” interativamente novos padrões musicais: elas manipulavam as peças coloridas até que escutassem algo que “fizesse sentido” para elas. Boa parte dos usuários preferiu descobrir como o sistema funcionava a receber instruções específicas. A colaboração surgia naturalmente, principalmente entre grupos de amigos e familiares.

De forma geral, as opiniões que recebemos dos participantes foram positivas. Apesar de termos apresentado um protótipo, os algoritmos se mostraram eficazes e robustos, embora, devido a condições adversas de iluminação e a características físicas do local da instalação (que faziam a mesa balançar de vez em quando), algumas recalibrações do sistema tenham sido necessárias.

## CONCLUSÃO

Neste trabalho descrevemos o desenvolvimento do “Tangible Beats”, uma interface tangível colaborativa para a criação, expressão e exploração de padrões musicais. O Tangible Beats permite que o usuário manipule objetos coloridos sobre uma superfície plana. Cada cor corresponde a um instrumento musical diferente. A superfície plana é dividida numa grade retangular. As colunas são varridas repetidamente, de modo a tocar o padrão musical codificado com as peças coloridas.

Diferentemente de outros sistemas sequenciadores para exploração musical de interfaces tangíveis que se utilizam de mesas interativas sensíveis a toque, o Tangible Beats utiliza uma solução flexível e de baixo custo baseada numa câmera RGB-D para detectar os objetos utilizados para interação. A informação visual, inclusive a linha de varredura, é projetada sobre a superfície com um projetor de vídeo comum.

Um protótipo do sistema foi construído utilizando uma câmera Kinect e um projetor de vídeo. Várias pessoas sem formação musical foram capazes de utilizar o protótipo sem receber instruções específicas, e pudemos verificar que ele promove e estimula a colaboração.

## Agradecimentos

Os autores agradecem à CAPES e à FAPESP (2011/00267-1), que financiaram o desenvolvimento deste projeto, bem como à equipe da Tenda Cultural Ortega y Gasset da USP pela valiosa assistência prestada durante a apresentação do protótipo.

## BIBLIOGRAFIA

1. Bradski, G., and Kaehler, A. *Learning OpenCV*. O’reilly, 2008.
2. Ishii, H., and Ullmer, B. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, ACM (1997), 234–241.
3. Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, ACM (2007), 139–146.
4. Kramer, J., Parker, M., Herrera, D., Burrus, N., and Echtler, F. *Hacking the Kinect*. Apress, 2012.
5. Lee, J.-Y., and Yoo, S. I. An elliptical boundary model for skin color detection. In *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*, Citeseer (2002).
6. Mann, Y., Lubow, J., and Freed, A. The tactus: a tangible, rhythmic grid interface using found-objects. In *NIME 2009* (Pittsburgh, PA, 03/06/2009 2009).
7. Marinho, B. N. S. F. C., Bergamo, M. L., and Mota, R. R. Programação para artistas. In *VII Brazilian Symposium on Games and Digital Entertainment (SBGames 2008)*, Belo Horizonte, MG–Brazil (2008).
8. Newton-Dunn, H., Nakano, H., and Gibson, J. Block jam: A tangible interface for interactive music. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*, NIME ’03, National University of Singapore (Singapore, Singapore, 2003), 170–177.
9. Wilson, A. D. Using a depth camera as a touch sensor. In *ACM international conference on interactive tabletops and surfaces*, ACM (2010), 69–72.
10. Xiao, R., Harrison, C., and Hudson, S. E. Worldkit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 879–888.

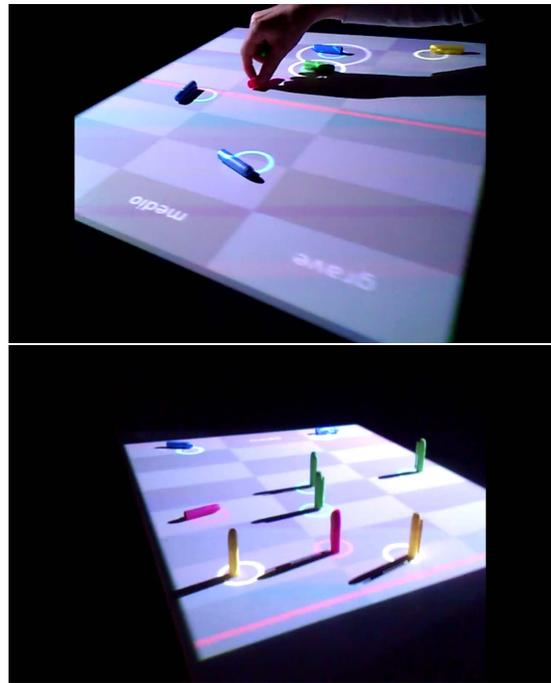


Figura 5: Tangible Beats na exibição de arte.