

LeinKlang : Novos Instrumentos de Expressão Sonora

Paulo Ferreira-Lopes^{†*}, Filipe Jensen[†]

[†]CITAR | Centro Investigação Ciências e Tecnologias das Artes – Universidade Católica Portuguesa

^{*}Z.K.M. | Zentrum für Kunst und Medientechnologie - Lorenzstr 19 , 76135 Karlsruhe – Germany

pflzkm.de fjensenlf@yahoo.co.uk

ABSTRACT — *Este artigo tem como objectivo principal descrever alguns aspectos de construção de um protótipo, apresentando genericamente alguns das problemáticas resultantes da realização do design do respectivo interface. A palavra LeinKlang (do alemão) tem como significado o ecrã (Leinwand) de sons (klang). Este projecto consiste na realização de um instrumento de música digital. O aspecto físico do instrumento apresenta-se sob forma de um ecrã, destinado a projectar sombra ou vídeos do movimento dos utilizadores, atrás do qual se encontra, intencionalmente dissimulados, um conjunto de 24 pequenos altifalantes.*

INDEX TERMS — *criação multimédia e audiovisual, interface design ; percepção ; som e visualização.*

I. INTRODUÇÃO

Os 24 altifalantes que constituem a “tela de projecção sonora do instrumento”, encontram-se homogeneamente distribuídos no rectângulo delimitado pelas margens do ecrã (3,5 x 2,5 metros). Este instrumento para além das características normais de um instrumento de música digital [3], permite também “visualizar o som”, a velocidade de deslocação do mesmo e o seu posicionamento espacial (numa primeira fase sob forma de coordenadas cartesianas – x,y). A velocidade das acções, bem como a distância entre o ecrã e o utilizador introduzem variáveis tanto no âmbito frequencial como no âmbito da dinâmica do som. As diferentes combinação destas variáveis permitem simultaneamente a manipulação espectral do som, como por exemplo associar a determinadas coordenadas regiões de filtragem de um espectro ou regiões de convolução de diferentes espectros. Tendo este projecto como objectivo principal a realização de uma trabalho sobre as modalidades de controle e de manipulação de instrumentos de musica digital através da interfaces visuais, não se poderia deixar de incluir algumas alusões com reflexões de ordem prática em torno da visualização do som [9] e dos modos de jogo associados ao binómio visual/auditivo. Para este efeito, propomos assim a realização de duas versões do instrumento, uma com suporte visual, e outra sem.

II. CONCEITO E DESCRIÇÃO GLOBAL PROJECTO

Num contexto global, poderíamos incluir esta proposta no domínio bastante abrangente dos jogos. Efectivamente como primeira abordagem, pretende-se que o contacto e a interacção entre o utilizador e o instrumento conduza a um espaço de entretenimento e logo também lúdico [7]. Numa abordagem mais avançada e num nível de conhecimento musical mais aprofundado pretende-se que o utilizador possa efectivamente compor som/música de forma consciente e com o domínio pleno das possibilidades vastíssimas do instrumento. Finalmente numa abordagem ainda mais complexa, pretende-se que o instrumento permita para além do aspecto lúdico do jogo e do aspecto estruturalista da composição a visualização do “sonoro”. Tratando este projecto da associação de três aspectos (jogo, composição visualização) com objectivo de estabelecer um universo caracterizado por uma consciência fictícia podemos afirmar que o mesmo introduz um carácter inovador relativamente ao no panorama global da construção de instrumentos de música digitais.

III. TIPOLOGIAS DE INTERACÇÃO COM A MATÉRIA SONORA

Fisicamente, o instrumento apresenta-se com a grelha de altifalantes, disposta numa matriz de em oito linhas verticais por três horizontais. Esta grelha pode-se apresentar escondida atrás de uma tela, onde é projectada a silhueta do utilizador em tempo real ou alternativamente, a grelha pode aparecer completamente “desnudada” permitindo assim visualizar parcialmente o “hardware” do instrumento : os altifalantes incrustados em placas de vidro acrílico transparente ladeados pelos respectivos amplificadores individuais do sinal áudio e respectiva cablagem. A possibilidade de apresentação dupla da instalação tem como base o estudo de dois diferentes tipos de interacção com o instrumento musical[3]. Assim por um lado, temos um instrumento que procura dirigir a interacção pelo lado da visualização do corpo do

interveniente, nas que as acções do utilizador são repercutidas em tempo real na tela. Por outro lado, a segunda variante pretende induzir o interveniente num âmbito de percepção acústica, onde os parâmetros musicais exigem uma audição mais acurada e consequentemente mais localizada. (ver Fig. 1)

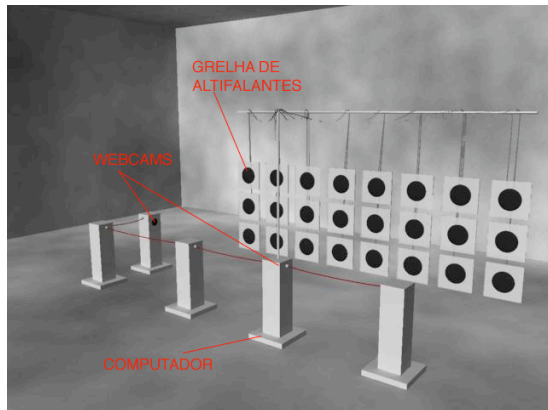


Fig. 1 - Instalação do instrumento com placas de acrílico

Duas *webcams*, dispostas em eixos opostos (no plano horizontal) seguem o movimento do utilizador, e estão escondidas dentro de duas colunas que circunscrevem o espaço destinado à interacção. O computador que processa o algoritmo de seguimento de vídeo e de geração de som encontra-se também na coluna central.

Na versão com opção de projecção de imagem, o utilizador tem ao seu dispor uma tela que espelha de forma mais ou menos deformada em frente da grelha de amplificadores (ver Fig. 2), o seu próprio movimento e a sua localização espacial. Estes dois aspectos irão desencadear o processo de interacção. Assim o movimento do utilizador delimita de forma análoga o numero de fontes de som activas e consequentemente a área de difusão sonora. O seu posicionamento espacial, aproximação à tela, determina a dinâmica de projecção do som de forma bipolar [3] : quanto mais o utilizador se afasta, mais o instrumento tenta compensar a dinâmica aumentando consequentemente o volume do som. Todavia este aspecto indicia também uma analogia à dimensão do espaço acústico das fontes : onde a volumetria arquitectural cresce em proporção com o volume sonoro. Nesta versão, as fontes de difusão sonoras que delimitam a figura do utilizador (os mini altifalantes) emitem normalmente sons gerados ou lido por um numero de instâncias sonoras reduzidas. De forma global a interacção pode ser caracterizada por uma metamorfose dum universo sonoro particular para um universo global [2].

No caso da versão sem tela, tendo em conta que se pretende induzir o utilizador num universo sonoro muito mais detalhado e de audição extremamente acurada, todas as fontes se encontram activas (os mini altifalantes) emitindo sons gerados ou lido até ao máximo de vinte e quatro instancias diferentes (numero máximo de altifalantes). Todavia tendo em vista uma abordagem do global para o particular[2], as coordenadas do perfil da cabeça do utilizador permitem que as fontes sonoras delimitadas por esta figura (quase uma circunferência) se excitam e modificam de forma mais variada e rica que as restantes. Desta forma é possível perceber uma espécie de paisagem sonora de fundo, perturbadas ou metamorfoseadas por elementos precisos e passíveis de uma composição sonora especifica e de acordo com a manipulação do utilizador.

O modulo de tratamento e geração de som foi programado em MAX/MSP, sendo particularmente o algoritmo de distribuição multicanal baseado na biblioteca SPAT do IRCAM.

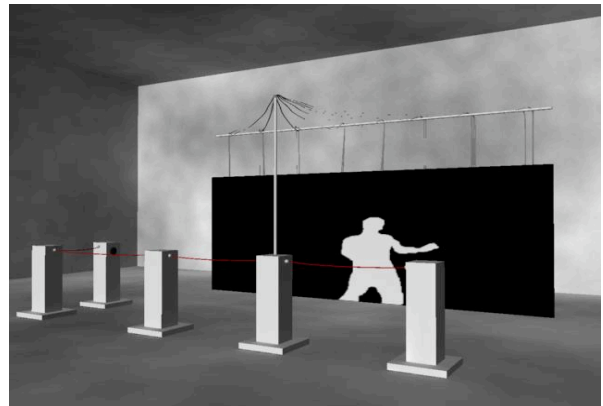


Fig. 2 - Projecção do em tempo real da silhueta do utilizador.

IV. MODULO DE SEGUIMENTO DE IMAGEM

O objectivo deste módulo consiste na elaboração do algoritmo que efectua o seguimento do utilizador face ao ecrã. O propósito inicial seria o seguimento de um máximo de sete pontos, correspondentes às extremidades do corpo humano (cabeça, ombros, mãos, joelhos), mais tarde, porém, decidiu-se que a quantidade de informação resultante da projecção sonora dos altifalantes seria redundante face ao utilizador. Este, estando próximo do ecrã, não teria a percepção do som que é projectado, por exemplo, na área respectiva aos seus joelhos, pelo que decidimos alterar o algoritmo de modo a localizar e seguir apenas a posição da cabeça, deixando todavia em aberto a

possibilidade de manter as restantes fontes sonoras em actividade de relevância inferior.

O método escolhido foi o fluxo óptico, nomeadamente o método de Lucas Kanade. Fluxo óptico define a análise da quantidade e direcção de movimento, normalmente traduzida em vectores entre pixeis [6][6]. O método de Lucas Kanade caracteriza-se pela sua robustez no que diz respeito à detecção de movimento nas extremidades de um objecto relativamente grande, ou seja, a variação do seu perímetro na matriz de vídeo, e a falta de capacidade de detectar movimentos rápidos.

No que concerne ao diferentes passos de programação foi escolhido o kit de programação o Max/MSP Jitter [10]e a biblioteca de Jitter, CV.Jit de Jean Marc Pelletier [11],que disponibiliza uma série de objectos baseados no método Lucas Kanade para a detecção e movimento de imagem.

Para isso usámos simultaneamente métodos de subtracção de fundo e “blob tracking”. A subtracção de fundo consiste na eliminação de toda informação visual não relevante de uma matriz de vídeo, ou seja, de tudo aquilo que não interfere no *tracking* que se pretende realizar. Neste projecto, tratou-se especificamente de subtrair à matriz toda a informação não relativa ao utilizador, nomeadamente, a tela e/ou a grelha de amplificadores, sendo a matriz resultante posteriormente é “binarizada”. Quando uma imagem é capturada pela câmara (qualquer dos objectos que não tenha sido subtraído), apresenta-se na matriz como uma mancha branca (uma “blob”) num fundo preto (*background* subtraído - ver Fig. 3.)

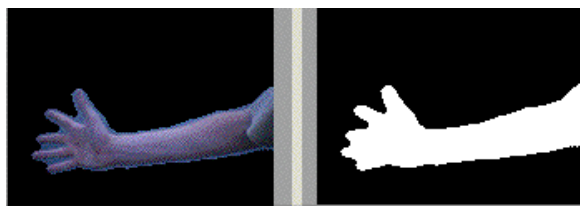


Fig. 3 - Subtracção de fundo e “blob” resultante da detecção de um braço

A partir do momento em que um “blob” é detectado, a biblioteca CV.Jit permite extrair várias informações sobre a área resultante, nomeadamente:

- Centro: coordenadas do centro da área coberta pelo blob;
- Orientação: Ângulo (em radianos) da disposição aproximada da área coberta pelo blob.;
- Rectângulo envolvente: Rectângulo resultante do cálculo da área descrita pelos pixeis extremos nos eixos de x e y da área do blob. (4 valores: x_1, y_1, x_2, y_2).

Através desta informação pudémos calcular as coordenadas da cabeça do “blob” que representa o utilizador na matriz de vídeo. A coordenada y é simplesmente o valor do pixel mais alto no eixo y da matriz de vídeo. Quanto à coordenada x, o processo de calculo não é directo, sendo a coordenada x, calculada através da orientação do “blob” resultante. Para tanto foi usada a formula que calcula coordenadas cartesianas a partir de um ângulo em radianos e um valor de raio, onde x e y são as coordenadas cartesianas, r é o valor do raio e α o ângulo (ver form. 1 e form. 2).

$$y=r*\cos\alpha$$

form. 1

$$x=r*\sin\alpha$$

form. 2

Tendo em conta que o ângulo é calculado pelo objecto cv.jit.orientation, resta apenas calcular o raio. Este ultimo corresponde á distância entre o centro da área do “blob” e o seu ponto mais alto (y). Seguidamente foi realizado um processo de escalonamento de valores resultante para o intervalo de valores da matriz de video capturada (0 a 119) e o conjunto de coordenadas já correspondia, com uma margem de erro desprezível, ao posicionamento da cabeça do utilizador. (ver Fig. 4)

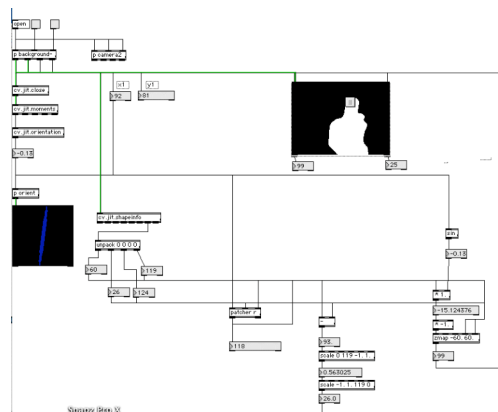


Fig. 4 O “patch” de “tracking” à posição da cabeça

IV. CONCLUSÃO

A construção de novos instrumentos musicais ou de expressão musical baseados no computador ou em

tecnologias digitais, encontra-se numa fase ainda incipiente devido às possibilidades que se abrem em diversos campos[4] [5]. Neste âmbito identificamos as tecnologias de tratamento do sinal vídeo em tempo real e as tecnologias de comunicação à distancia, particularmente a internet e os espaço da WEB [1]. No caso do protótipo que apresentamos tentamos integrar a tecnologias de tratamento e de geração de sinal vídeo em tempo real essencialmente com o objectivos de explorar, ainda que timidamente, um aspecto que nos parece determinante na percepção da matéria sonora ; a visualização. No caso da tecnologias de comunicação, este aspecto torna-se cada vez mais imperativo numa perspectiva da partilha publica do resultado de objectos que sob o ponto de vista geográfico incorporam uma certa invisibilidade na produção da arte ou de expressões artísticas. Este será sem dúvida o próximo passo que no conjunto das nossas reflexões iremos abordar e clarificar. Resta ainda acrescentar que alguma da timidez global, no que concerne a taxinomia e classificação [8] deste tipo de instrumentos exigem no futuro uma reflexão no sentido de ajudar a enquadrar de forma mais clara a natureza do instrumento e a sua ou suas áreas de acção.

AGRADECIMENTOS

Este trabalho foi realizado com o apoio do CITAR – Centro Investigação Ciências e Tecnologias das Artes da Universidade Católica Portuguesa, pela Fundação para a Ciência e Tecnologia de Lisboa e pelo ZKM | Zentrum für Kunst und Medientechnologie - Institut für Musik und Akustik de Karlsruhe. Agradecimento ao Álvaro Barbosa e António de Sousa Dias pela revisão deste trabalho.

REFERÊNCIAS

- [1] **BARBOSA, A.** ; *Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation* in Leonardo Music Journal (13), 2003
- [2] **FERREIRA-LOPES, P. And SOUSA DIAS, A** ; “Musique et int eraction: aboutissement, mutations te m etaphores de l'instrument de musique num erique” in actes des Journ ees d'Informatique Musicale. Paris: CCIM. ; Paris 2005.
- [3] **FERREIRA-LOPES, P.** ; * tude de mod eles interactifs et d'interfaces de contr ole en temps r eel pour la composition musicale*. Th ese de Doctorat ; Paris ; Universit e de Saint Denis - Paris VIII - D ep. de Sciences et Technologies des Arts, 2004.
- [4] **JORDA, S.** ; “*Digital Instruments and Players: Part I – Efficiency and Apprenticeship*” in Proceedings of 2004 International Conference on New Interfaces for Musical Expression NIME04 , Hamamatsu, Japan.
- [5] **JORDA, S.** ; “*Multi-user Instruments: Models, Examples and Promises*”, in Proceedings of 2005 International Conference on New Interfaces for Musical Expression NIME05 , Vancouver, BC, Canada, 2005
- [6] **LUCAS, B. D. and KANADE T.** ; “*An iterative image registration technique with an application to stereo vision*” in Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver, pp. 674--679. 1981
- [7] **MACHOVER, T.** ; “*Instruments, Interactivity, and Inevitability*” in Proceedings Proceedings of 2002 International Conference on New Interfaces for Musical Expression NIME02, Dublin , Ireland, 2002.
- [8] **MARSHALL, M.T. and WANDERLEY, M.**; “A survey of sensor use in digital musical instruments” [http://www.music.mcgill.ca/musictech/idmil/projects/Sensor Survey](http://www.music.mcgill.ca/musictech/idmil/projects/SensorSurvey)
- [9] **SEDES, COURRIBET, THIEBAUT, VERFAILLE** ; “*Visualisation du sonore, vers la notion de transduction : une approche en temps r eel*, in Espaces Sonores - Actes de Recherche Anne Sedes,  ditions musicales Transatlantiques, Paris, 2003.
- [10] <http://www.cycling74.com>
- [11] <http://www.iamas.ac.jp/~jovan02/cv/>