

Visualização e Exploração Interactiva de Espaços de Vídeo através de Cor e Movimento

João Martinho
LaSIGE, Faculdade de Ciências
Universidade de Lisboa
jdmartinho@gmail.com

Teresa Chambel
LaSIGE, Faculdade de Ciências
Universidade de Lisboa
tc@di.fc.ul.pt

Abstract

Temos assistido a um rápido crescimento na criação e disponibilização de vídeos em espaços de informação de acesso generalizado. O próprio vídeo é um meio bastante rico, incluindo imagens, texto e áudio que variam ao longo do tempo. Esta riqueza contribui para que estes espaços de informação sejam bastante interessantes de explorar, permitindo a comunicação de uma grande quantidade de informação e constituindo uma excelente plataforma para expressar criatividade. No entanto, inerente a esta riqueza, existe uma complexidade significativa. As técnicas de visualização podem auxiliar a lidar com esta complexidade e a expressar e explorar a riqueza destes espaços de informação.

Este artigo apresenta um ambiente interactivo para visualizar e explorar espaços de vídeo, com um foco em aspectos culturais e salientando propriedades tais como a predominância de cor, ritmo e movimento, tanto ao nível do espaço de vídeos como ao nível dos vídeos individuais. Permite capturar, experienciar e explorar propriedades e relações entre os vídeos, oferecendo a possibilidade de obter novos olhares sobre a cultura e de influenciar a sua expressão de forma criativa.

Keywords

Espaços de Vídeo, Visualização, Análise de Imagem e Vídeo, Criatividade, Cor, Movimento, Arte Digital

1. INTRODUÇÃO

O vídeo é actualmente um meio amplamente usado no dia a dia. Com a propagação das câmeras digitais, telemóveis, redes de banda larga e websites de partilha de vídeos, é bastante comum aceder a vídeo digital em cenários públicos e privados. O vídeo é também um meio bastante rico que combina imagem e som, oferecendo grandes quantidades de informação que pode ser usada como plataforma para explorar a criatividade. Os pioneiros da Vídeo Arte, por exemplo, têm vindo a explorar esta criatividade, usando arte de instalação e *performance* juntamente com vídeo de forma a criar experiências imersivas para o espectador. No entanto, esta riqueza que faz com que os espaços de vídeo sejam interessantes, é ao mesmo tempo bastante complexa de lidar.

Um dos problemas prende-se com o facto do vídeo ser um meio de informação não estruturada, tornando o seu acesso difícil. Descritores semânticos podem ser usados para anotar alguma informação do vídeo: por parte dos utilizadores, uma técnica comum por exemplo em repositórios de vídeo *online* como o YouTube; ou através de segmentação de informação. Dados de baixo nível, como a cor e informação da filmagem, duração do vídeo e informação de cena, oferecem informação adicional que pode ser obtida através de análise de vídeo. Uma vez que esta informação tenha sido extraída podemos usá-la em nosso proveito, para organizar espaços colectivos de

vídeo, fazer pesquisas, ajudar em processos de edição e também para oferecer novas formas de visualização e interacção.

As técnicas de visualização podem ajudar a expressar a riqueza destes espaços de informação. A visualização de vídeos pode ser uma forma eficaz e intuitiva de mostrar informação importante do vídeo. No entanto, ainda existe uma falta de técnicas eficazes capazes de apresentar informação relevante e complexa através de processamento automático de vídeo. Uma alternativa é disponibilizar uma vista global, extraindo informação interessante e apresentando-a de uma forma significativa [Daniel03]. As técnicas de sumarização, por exemplo, permitem visualizar vídeos de formas mais simples e concisas, baseando-se em propriedades como o movimento, ritmo e mudanças de cena. Ainda no movimento da Vídeo Arte, alguns dos trabalhos fazem uso da visualização como uma ferramenta para passar uma mensagem ou um significado ao espectador.

Avanços na área da visualização de dados têm surgido de investigações baseadas principalmente nas áreas da percepção visual e da cognição. A tendência vai no sentido da evolução para cenários onde as ferramentas se tornam invisíveis enquanto os utilizadores se tornam imersos na exploração da informação [Few07].

Este artigo apresenta um ambiente de suporte à visualização e exploração interactiva e criativa de vídeos,

com um forte ênfase nas propriedades de cor e movimento, cruzando o acesso a informação, cultura e arte digital. O ColorsInMotion (Figura 1) oferece meios de análise de vídeo, pré-existente ou capturado do mundo real através de uma câmara, e mecanismos de visualização e navegação em espaços de vídeo, evidenciando propriedades como a predominância de cor, ritmo e movimento. A interacção pode ser feita através de meios convencionais como monitor, rato e teclado, ou através de ecrãs sensíveis ao toque. Estamos também a desenvolver formas mais naturais de interacção, sem recurso a dispositivos, especialmente adequadas para instalações e contextos de interacção ambiental.

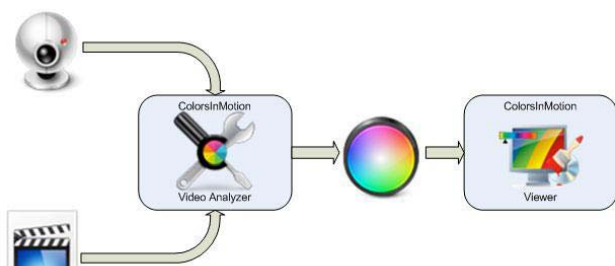


Figura 1: Arquitectura do sistema ColorsInMotion

Na secção 2 apresentam-se os trabalhos relacionados mais relevantes. A secção 3 apresenta a aplicação de visualização do ColorsInMotion. A secção 4 descreve o VideoAnalyzer, o módulo responsável por processar e analisar vídeos para o sistema, e na secção 5 o artigo termina com conclusões e a identificação de perspectivas para trabalho futuro.

2. TRABALHO RELACIONADO

Existe um conjunto de ferramentas e aplicações de visualização. Encontra-se uma selecção bastante significativa em [Perez08]. A maioria é usada para visualizar redes sociais, representações da Internet, música, e alguns websites como o Amazon e o Flickr. Algumas são criadas para fins artísticos, enquanto outras são vocacionadas para a exploração e análise de dados. No entanto, não existem muitas que façam visualização de vídeo. Algumas excepções incluem a visualização de vídeos no YouTube e o VideoSphere. A partir de cada vídeo do YouTube, o utilizador pode aceder a uma vista em 2D que representa os vídeos como imagens estáticas circulares espalhadas no ecrã e rodeados pelos vídeos relacionados, dando acesso às páginas tradicionais de visionamento de cada vídeo. Permite uma navegação baseada em vídeos relacionados, mas disponibiliza informação e funcionalidade limitada sobre os vídeos ao nível do espaço de vídeos. O VideoSphere [Bestiario08] representa os vídeos em volta de uma esfera tridimensional, com ligações entre os vídeos, reflectindo ligações semânticas, mas não oferece qualquer forma especial de visualização dos vídeos individualmente, apenas as tradicionais representações por keyframe estática, como no YouTube, e o visionamento normal do filme.

Existe um número considerável de trabalhos que explora novas formas de ver um vídeo individualmente. Em [Fels99], Fels e Mase mostram uma vista tridimensional do vídeo considerando os dados do vídeo como um volume, onde a terceira dimensão é o tempo, similar ao que Daniel e Chen apresentaram em [Daniel03]. Este tipo de visualização oferece ao utilizador uma maneira de ver a evolução de uma determinada área do vídeo ao longo da sua duração, usando os eixos x e y para a largura e altura e o eixo z para representar o tempo, dando profundidade ao volume. As técnicas de slit scan são também capazes de capturar acções temporais em imagens estáticas, tornando-os candidatos interessante para a representação de vídeos [Levin08]. Em [Nunes07], o Timeline adopta slit scans para permitir uma fácil e rápida exploração do historial dos vídeos capturados pelas *webcams* para visualizar acção e presença de participantes em cenários colaborativos.

Em [Boreczky00] os autores fazem uso de uma representação baseada em banda desenhada de forma a sumarizar um vídeo individual. As frames são agrupadas usando a técnica de agrupamento hierárquico aglomerativo, de forma a produzir segmentos que podem ser representados por uma keyframe, que é mostrada. Comparável a esta visualização, em [Irani95], os autores fazem uso de uma representação baseada em mosaicos para apresentar frames significativas numa sequência de vídeo. Esta representação pode ser usada para fazer um sumário visual do vídeo. Em trabalhos anteriores exploramos este e outros tipos de representação para indexar vídeo e auxiliar a navegação em hipervídeo [Chambel02]. Outros trabalhos relacionados, como [Finkelstein98] ou [Klein02], usam técnicas de pesquisa de imagens com base no conteúdo para fazer o processo de selecção necessário para criar mosaicos de vídeo. Em [Martin04], o autor usa algumas destas técnicas para tentar recriar um vídeo inteiro com mosaicos. Os trabalhos baseados em mosaicos fazem uso de histogramas de cor e diversas técnicas de análise de imagem de forma a encontrar as imagens que mais se adequam as regiões a substituir no mosaico.

Num trabalho prévio [Rocha08], oferecemos uma visualização interactiva em 3D juntamente com navegação de vídeo, como uma instalação artística, para explorar propriedades culturais e ligações entre diferentes vídeos tanto a um nível global onde uma colecção de vídeos é representada por loops organizados por categorias semânticas como países, temas e autores; e a um nível individual onde são usadas propriedades do vídeo como a cor e o brilho para fazer uma representação tridimensional do vídeo com a qual o utilizador pode interagir enquanto ele é apresentado.

Outros trabalhos que nos mostram diferentes formas de navegar numa colecção de vídeos, desta vez com o objectivo de ajudar no processo de edição de vídeo, incluem [Girgensohn00], onde o sistema de edição Hitchcock cria pilhas de vídeos organizadas por semelhança de cores, oferecendo uma maneira de organizar visualmente uma grande colecção de vídeos.

De forma a ter dados suficientes para efectuar este tipo de organização, também existem alguns trabalhos que se focam na parte da análise de imagem e dados contidos no vídeo, como em [Kerminen99], onde os autores criaram um sistema de recuperação de imagens baseado em histogramas de cor. Para criar os histogramas de cor analisam as imagens pixel a pixel e quantizam o espaço de cor. Em [Saykol04] os autores também usam uma técnica similar para pesquisa de imagens por cor, forma e textura. O Multicolr Search Lab [Idée08] encontra imagens baseadas em cores seleccionadas. Os utilizadores podem escolher até dez cores, e podem repetir aquelas que querem que sejam predominantes. As imagens são seleccionadas e pré-processadas, a partir de uma base de dados de fotografias do Flickr e do Alamy Stock Photography. Este tipo de websites de partilha de fotos normalmente apenas permite que se façam pesquisas semânticas usando palavras chave e anotações, mas nunca cores. O mesmo é válido para websites de partilha de vídeo como o YouTube.

A criatividade também é um aspecto importante no nosso trabalho. Existem vários projectos que usam formas originais e diferentes de visualizar e organizar vídeos e outros tipos de informação. Por exemplo, em [Sims91], Sims usou algoritmos genéticos e a teoria da evolução para criar gráficos. Numa mesma perspectiva, também explorámos algoritmos evolucionários em trabalhos prévios [Chambel07]. Novas sequências de vídeo eram combinadas e seleccionadas com base nas suas características representadas como anotações dos vídeos, quer pela definição de critérios quer efectuando selecções ao longo da evolução da população de vídeos. Um dos objectivos incluía explorar e descobrir relações culturais e arquitectónicas em vídeos de Portugal e do Brasil. O Soft Cinema [Manovich] é uma instalação dinâmica baseada em narrativas ambientais como forma de representar a experiência subjectiva de uma pessoa a viver numa sociedade de informação global. É apresentada uma série infinita de filmes narrativos construídos na altura, usando regras definidas pelos autores, e uma base de dados de media contendo algumas horas de vídeo, animação, narração por voz e música. Na área do Video Jockeying (VJing ou VJ) [Makela06], os artistas criam arte visual em movimento, normalmente usando vídeos em grandes ecrãs ou telas, em eventos como concertos, discotecas, e por vezes juntamente com outra arte de *performance*. Ferramentas de edição e visualização criativa de vídeo podem ser usadas como suporte para estes artistas.

O processo criativo normalmente tem um certo grau de aleatoriedade e incerteza – como Pepperell em [Pepperell02] refere, “a criatividade é a resposta do organismo à impossibilidade de prever o ambiente”. Levando esta ideia em conta, muitas das técnicas mencionadas tentam aplicar esta metáfora do sistema dinâmico e evolucionário a colecções de vídeos ou gráficos, combinando as suas propriedades e regras intrínsecas com o contexto e a interacção do utilizador para induzir e suportar o processo criativo de formas inovadoras.

3. COLORS IN MOTION

ColorsInMotion é uma aplicação interactiva para visualizar e explorar vídeos, baseando-se principalmente nas propriedades de cor e movimentos. No espaço de vídeos, o utilizador pode pesquisar, ver, comparar e interagir com uma selecção de vídeos através de diferentes perspectivas, com ênfase em diferentes propriedades e criando diferentes vistas. Neste nível, o utilizador pode seleccionar e navegar para cada um dos vídeos, onde os pode visionar e explorar as suas propriedades individuais. Nesta secção vamos introduzir os conceitos principais e as diferentes opções para interagir no ColorsInMotion.

O espaço de vídeos é concebido como um sistema físico de partículas que se movem no ecrã de acordo com a sua semelhança de cor e a interacção com o utilizador, permitindo fazer pesquisas, comparações, navegar e visualizar de uma forma criativa.

3.1 Vistas de Cor e Movimento no Espaço de Vídeos

Ao nível do espaço de vídeos, o utilizador pode mudar entre diferentes vistas para obter perspectivas sobre o conteúdo dos vídeos e as suas propriedades de cor e movimento.

Loops de vídeo tradicionais incluem frames extraídas em intervalos constantes de tempo e apresentadas em ciclos infinitos - dão uma perspectiva sumariada do conteúdo (Figura 2¹). Loops por cor média e por cor dominante, apresentam frames seleccionadas com base na semelhança com estas propriedades de cor, oferecendo uma perspectiva sumariada mais representativa das cores do vídeo.

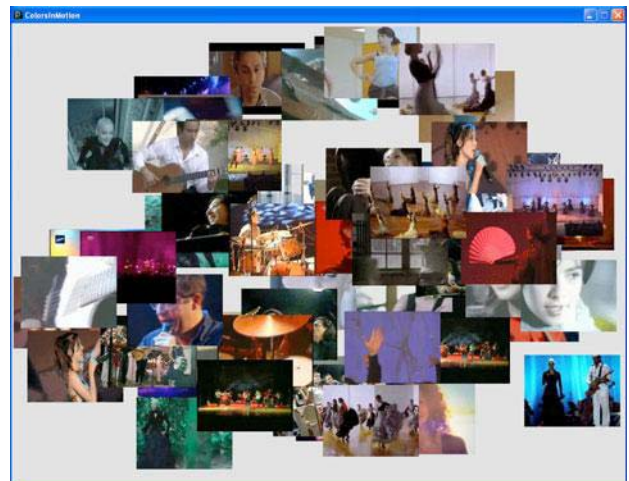


Figura 2: Loops de vídeo no espaço de vídeos.

¹ A cor de fundo do ColorsinMotion é configurável. Para efeitos de impressões com o melhor contraste possível, foi escolhida uma cor clara. Numa instalação, um fundo escuro em ecrã inteiro pode ser mais adequado para evidenciar os vídeos.

Para dar uma relevância maior às cores, os vídeos podem ser representados por círculos preenchidos com a cor média ou dominante, ou também podem ser representados por rectângulos preenchidos por tiras das cores dominantes do vídeo, onde a largura de cada tira é proporcional ao número de frames onde domina no vídeo. Estas vistas são apresentadas nas Figuras 3 e 4.



Figura 3: Evolução do sistema de partículas no espaço de vídeos - vistas da cor média dos vídeos.

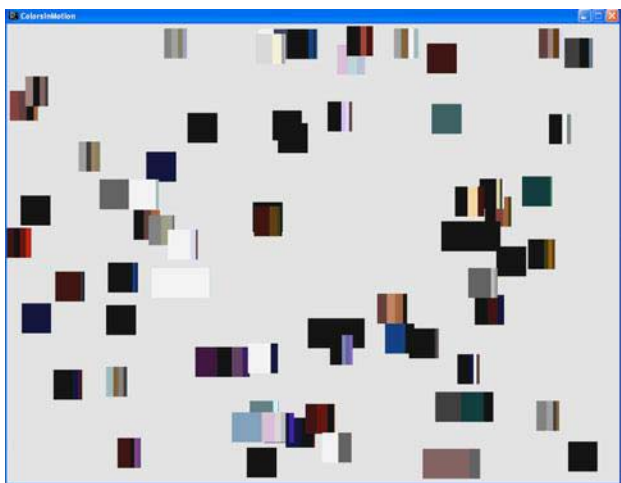


Figura 4: Vistas de cores dominantes no espaço de vídeos.

Para evidenciar aspectos do movimento, os vídeos podem ser representados por loops de cenas ou por slit scans. Nestes loops (Figura 5), cada imagem captura o movimento através da média das cores dos píxeis em cada posição da imagem ao longo do tempo. Os slit scans [Levin08] capturam, numa imagem e em sequência, a acção que toma lugar na secção central das frames do vídeo. Em vez de criar uma imagem desfocada e manchada como acontece nos loops, aqui a acção é apresentada em sequência produzindo uma imagem com uma largura que representa a duração do vídeo. Nesta vista, as mudanças de cena podem ser vistas pelo utilizador através das mudanças abruptas que ocorrem horizontalmente na imagem.

No espaço de vídeos (Figura 6), os slit scans são apresentados de forma condensada, mas podem ser expandidos através da acção do utilizador ao colocar o cursor por cima da imagem respectiva, como está exemplificado, ao lado, para um vídeo de danças Sevilhanas em tons púrpura.

Todas estas vistas são criadas usando os dados extraídos durante a fase de análise de vídeo, descrita na secção 4.

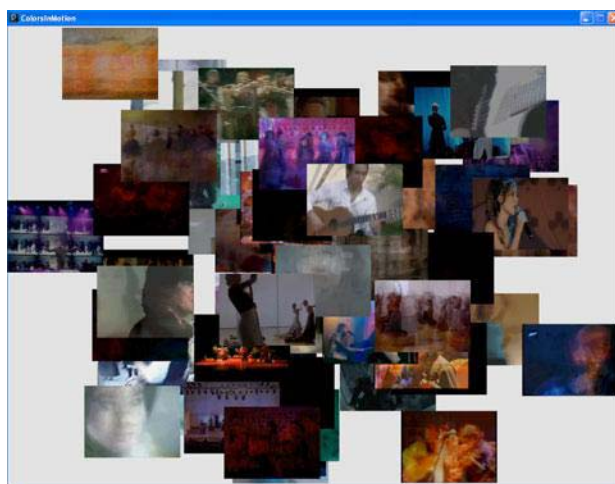


Figura 5: Loops de cenas no espaço de vídeos.

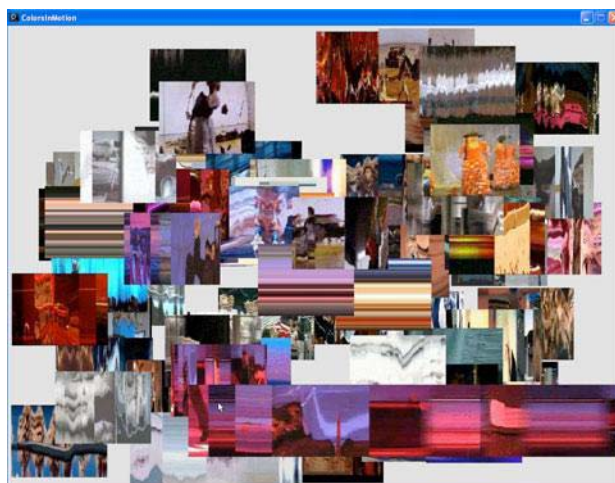


Figura 6: Slit scans no espaço de vídeos.

3.2 Ao Nível dos Vídeos Individuais

A partir do espaço de vídeos podemos seleccionar e aceder a um vídeo individual. Esta vista permite visionar o vídeo, aceder a detalhes e às diferentes vistas do vídeo. A Figura 7 mostra uma vista individual do vídeo de danças Sevilhanas. O vídeo é mostrado no centro. Na fila de cima, os aspectos de cor são evidenciados através do loop por cor média, o círculo preenchido por cor média (púrpura), seguido do rectângulo que mostra as cores dominantes, o círculo preenchido com a cor dominante (vermelho acastanhado) e o loop por cor dominante. Note-se que a cor dominante prevalece no rectângulo como seria de esperar. À esquerda do vídeo temos o loop tradicional de vídeo e à direita o loop das cenas,

evidenciando aspectos do movimento, complementados através da imagem em slitscan apresentada em baixo.

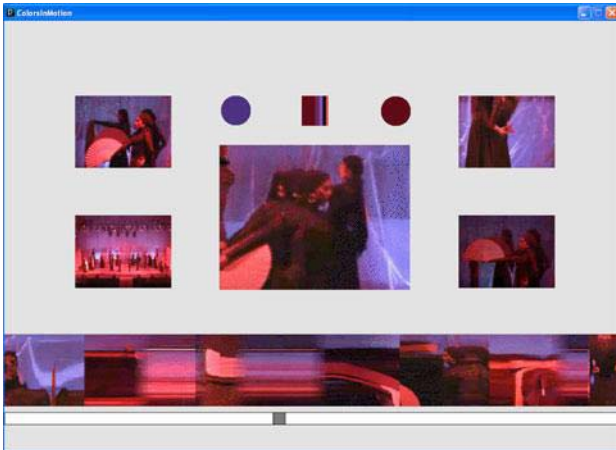


Figura 7: Vista de um vídeo individual.

Nesta vista, os utilizadores acedem às perspectivas complementares sobre o vídeo, e podem a partir daqui aceder a todas as vistas do espaço de vídeos (seleccionando a perspectiva desejada), onde podem comparar este vídeo com os restantes nesta perspectiva escolhida. Pode-se por exemplo perceber as diferentes cores e ritmos em vídeos de dança de diferentes autores e países na vista de slit scans.

3.3 Pesquisas no Espaço de Vídeos

No espaço de vídeos demonstrado apresentamos uma selecção de vídeos num contexto cultural, com temas como a música e a dança de diferentes autores e países. O utilizador pode pesquisar e seleccionar vídeos através da cor média ou dominante, seleccionando uma ou mais cores de uma paleta, ou do mundo real através de uma câmara (Figura 8). Poderá escolher mais de uma cor, especificando percentagens de predominância para cada uma. Em ambos os casos pode ser definido um limiar de proximidade para controlar a precisão da pesquisa. Os resultados são apresentados na mesma vista em que a pesquisa foi feita.



Figura 8: Interface de pesquisa por cor.

As vistas do espaço de vídeos são criadas a partir de partículas que deslizam no ecrã. As forças entre os vídeos dependem da sua cor relativa, atraíndo-se e repelindo-se conforme a semelhança. Desta forma, estas vistas permitem por si só um tipo de pesquisa com base na semelhança de cores, quer apenas observando a evolução do sistema à medida que as partículas semelhantes se juntam, quer interagindo com ele, reposicionando partículas e alterando-lhes a velocidade, para alterar o foco e influenciar o seu caminho. No início, as partículas começam em localizações aleatórias e vão evoluindo à medida que as forças actuam. Na Figura 2 podemos ver o

sistema no início. Na Figura 3, os círculos já se moveram durante algum tempo, atraíndo os de cor semelhante.

Existe ainda uma vista do espaço de vídeos que só pode ser obtida através de uma pesquisa específica – o mosaico de vídeos. Os mosaicos são obtidos através de imagens em camadas, compostas por elementos individualmente reconhecíveis que quando combinados sugerem uma imagem diferente no seu todo [Martin04]. Esta vista é composta de pequenos loops de vídeos de cor dominante que melhor se adequam à cor de determinadas regiões de uma imagem introduzida pelo utilizador para fazer a pesquisa. A imagem tanto pode ser arrastada para a aplicação, como pode ser capturada do mundo real com uma câmara. Os métodos de pesquisa usados para criar os mosaicos são idênticos aos que são usados para seleccionar os vídeos no espaço de vídeos com base em cores dominantes.

No futuro, através de reconhecimento de conteúdo em imagens, ou apenas com acesso a metadados, por exemplo, nos casos em que a imagem é anotada e colocada num repositório de imagens, a pesquisa pode ser reforçada com propriedades semânticas. Por exemplo, se o utilizador apresentar uma imagem da Torre Eiffel, todos os vídeos usados para criar o mosaico da torre incluiriam também a torre Eiffel, ou dependendo do âmbito, seriam de Paris ou de França. No contexto cultural da nossa demonstração, por exemplo, seleccionaria vídeos de música ou dança franceses.

Os mosaicos de vídeo podem conter múltiplas camadas de imagens perceptíveis e interessantes para usos comerciais, culturais e artísticos e para expressar a criatividade com auxílio de um processo de pesquisa automatizada.

3.4 Criatividade Interactiva com Cor em Movimento

O acaso é muitas vezes considerado um elemento importante no processo criativo e pode ser visto como um complemento ao controlo do utilizador na criação de obras de arte digitais [Kneller65, Paul03]. No ColorsInMotion, os utilizadores podem basear-se no nível de incerteza inerente às propriedades dos vídeos e especialmente à forma como as suas cores geram forças e fazem com que o espaço se organize, ou podem interferir seleccionando a vista a ser apresentada, os vídeos que são apresentados, e influenciando os movimentos e *zooming* para criar o resultado visual. Podem também basear-se na incerteza inerente a cada vídeo com o seu conteúdo particular, e controlar parâmetros como os limiares de proximidade para influenciar a estética na criação das vistas que capturam cor e movimento, como os loops de cenas, mosaicos de vídeo e slit scans.

No espaço de vídeos, definido com um mundo físico, todas as partículas têm a mesma massa e velocidade aleatória no início da evolução. Uma vez colocadas as partículas, são criadas forças entre todas elas. Como mencionado acima, estas forças podem ser de atracção ou repulsão, dependendo da semelhança de cores entre elas. Se a semelhança for alta, as partículas atraem-se, se as

cores forem muito diferentes irão repelir-se, evoluindo para uma organização por semelhança de cor. Assim, as cores semelhantes irão formar grupos, enquanto que as cores diferentes irão fugir desses grupos e colidir nas fronteiras do ecrã e voltar, perdendo alguma velocidade. Uma vez que não existe gravidade, as partículas podem deslizar livremente em qualquer direcção. A interacção entre elas eventualmente chegará a um ponto de equilíbrio, devido a uma pequena força de atrito introduzida no sistema e à tendência que têm em criar as distâncias relativas de acordo com as forças de atracção e repulsão. Desta forma, o movimento não é perpétuo.

Os utilizadores podem interagir com o sistema arrastando as partículas, influenciando a sua velocidade e direcção, e depois ver o resultado produzido. Para além de poder controlar o *zoom* da vista, é possível ainda criar um efeito de arrasto que as partículas vão deixando no ecrã. Desta forma, com uma mistura de aleatoriedade e controlo do utilizador, é possível criar efeitos visuais como o que se exemplifica na Figura 9, onde se usou a vista dos círculos por cor média com arrasto para criar o que se assemelha a uma pintura abstracta. Note-se como as cores semelhantes já se estão a aproximar, e que os utilizadores podem interferir com esta tendência em qualquer altura.

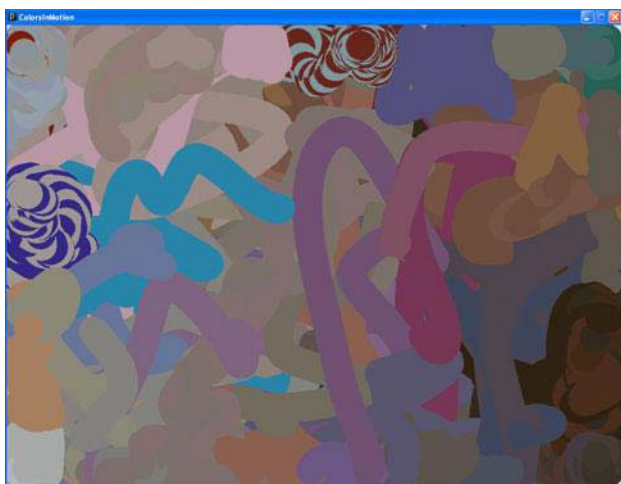


Figura 9: Rasto deixado pelas partículas na vista de cor média no espaço de vídeos.

Os utilizadores podem basear-se na incerteza inerente ao conteúdo individual e opcionalmente controlar parâmetros que podem influenciar a estética na criação de algumas das vistas.

A Figura 10 apresenta imagens conseguidas a partir da média dos píxeis em cada posição da imagem num vídeo de música, à esquerda, e num vídeo de dança, à direita. Estas imagens têm uma estética impressionista e capturam a cor e o movimento do vídeo. Alguns parâmetros podem ser controlados para influenciar a criação destas imagens. Estas imagens também podem ser vistas nos loops de cenas, na Figura 5.



Figura 10: Movimento em cenas de música e dança.

Diferentes parâmetros podem também influenciar a escolha de loops para os mosaicos de vídeo e a selecção de tiras nas imagens de slit scans, influenciando a criação do resultado visual.

3.5 Interacção com Cores e Movimento

Toda a interacção pode ser feita através do rato e teclado, e a maioria das acções pode ser feita através das acções de apontar, seleccionar e arrastar – até os valores podem ser introduzidos através de elevadores – o que também é adequado numa interface sensível ao toque. Para além da captura de cores e imagens do mundo real, para pesquisas por cores e criação de mosaicos de vídeo, também foram feitos alguns mapeamentos para interacção baseada na detecção de cores pré-definidas em regiões específicas exibidas em frente ao ecrã. Está ainda em desenvolvimento o mapeamento para uma interface baseada em gestos. Estas modalidades baseadas na cor e movimento podem oferecer uma experiência mais natural num ambiente de instalação, fazendo também a ponte entre a maneira como se interage e a informação que se está a ver, ambos baseados em cor e movimento.

A interacção com teclado e rato pode ser adequada para a maioria das situações. O teclado em particular pode ser o dispositivo preferido por utilizadores experientes, e por exemplo em ambientes de VJ, oferecendo um desempenho mais elevado e maior precisão na selecção das acções no momento certo. As modalidades alternativas de interacção providenciam flexibilidade para acomodar diferentes tipos de utilizador e cenários de utilização, no sentido de experiências mais naturais e potencialmente mais imersivas.

4. ANÁLISE DE VÍDEOS

O ColorsinMotion usa uma base de vídeos pré-processados. O VideoAnalyzer é o módulo responsável pela análise dos ficheiros de vídeo, pela produção de metadados e por algumas das visualizações dos vídeos. Os vídeos podem ser acedidos a partir de um repositório público como o YouTube, uma localização dada pelo utilizador ou capturados do mundo real a partir de uma câmara, oferecendo a hipótese de criar vídeos na hora. Esta secção apresenta as principais funcionalidades do VideoAnalyzer.

4.1 VideoAnalyzer e Vistas de Vídeo

O VideoAnalyzer processa e mostra o vídeo em tempo real (primeira imagem no canto superior esquerdo da Figura 11), mostrando a evolução da cor média e da cor dominante em cada frame numa visualização ao estilo de uma linha temporal (primeiras duas linhas horizontais na Figura 11).

Para analisar a cor do vídeos, são usados histogramas de cor. Os histogramas de cor são criados com um número de contentores especificados pelo utilizador. Quanto mais contentores tiver, mais cores individuais iremos capturar, no entanto à custa de algum desempenho. Os histogramas são usados então para obter a cor média do vídeo, as cores dominantes de todo o vídeo, as cores dominantes em cada frame e a percentagem de predominância destas cores (quadrado púrpura e gráficos seguintes na Figura 11). Estas são algumas das métricas que são usadas para indexar a colecção de vídeos, bem como para produzir algumas das visualizações.

A última fita na figura representa a visualização em slit scan que vai sendo construída também em tempo real durante a análise do vídeo.



Figura 11: Interface do VideoAnalyzer.

Para a visualização de cenas é capturada a evolução das frames do vídeo, acumulando a cor média em cada píxel, nas respectivas posições ao longo do tempo (imagem por baixo do vídeo no canto superior esquerdo da Figura 11). Se o vídeo tiver muito movimento, a imagem gerada sofrerá muito desfoque e ficará bastante escura. Caso contrário, uma imagem com algum movimento e com as impressões ao longo do tempo irá ser criada. Usando um limiar de proximidade entre cores, é feita uma análise frame a frame numa forma simples de detecção de cenas. Cada vez que é detectada uma mudança, a imagem que capturou o movimento da cena é adicionada ao loop de cenas, e é iniciada uma nova imagem para a cena seguinte.

Os loops tradicionais do vídeo apenas têm em conta a duração do vídeo. São criados através da extracção de um número de imagens desejado com intervalos regulares entre si.

Após a análise do vídeo e da apresentação dos resultados, podemos também extrair loops baseados nas medidas de cor. Estes incluem os loops por cor média, loops por cor dominante e ainda loops por uma cor à escolha do utilizador.

Os loops por cor média são criados comparando a cor média de cada frame com a cor média resultante da análise ao filme todo, e é possível controlar um limiar de proximidade entre as cores de forma a influenciar a precisão dos resultados. Também podemos escolher o número máximo de imagens a extrair para o loop. Os loops por cor dominante são semelhantes aos de cor média e são criados comparando o histograma de cor de cada frame com o histograma de cor do vídeo. Também é possível controlar a proximidade da cor e o número de imagens a extrair. O utilizador pode ainda criar um loop a partir de uma cor a sua escolha, que pode ser escolhida nas visualizações e gráficos apresentados ou ainda numa paleta de cores. Usa o mesmo método de extracção que os loops por cor dominante.

Todos os loops são gravados e relacionados com o respectivo vídeo para uso posterior. Também é gravada informação de metadados que sejam extraídos do vídeo para uso na navegação e pesquisa das colecções de vídeos. As visualizações dos círculos por cor média ou dominante, por exemplo, são criadas no módulo de visualização do ColorsInMotion, com base nestes dados.

4.2 Métricas e Espaços de Cor

Para a análise dos vídeos foram experimentados dois espaços de cor: RGB e HSB. No espaço de cor RGB, foi usada uma quantização uniforme, i.e. o número de contentores (especificado pelo utilizador) é o mesmo para cada uma das componentes, contribuindo todas com o mesmo peso para a percepção da cor. Este é um método comum de lidar com cores, mas não enquadra da melhor forma alguns aspectos perceptuais, podendo não corresponder completamente às expectativas dos utilizadores.

A quantização de cor do espaço HSB, por outro lado, não é uniforme mas leva a percepção humana em conta. Foca-se nas dimensões da cor (hue), saturação e brilho – facilmente entendidas pelos utilizadores – e permite ter em conta o facto de que os humanos são mais perceptíveis ao brilho que à cor. Assim, cores pouco saturadas podem ser percebidas pelo seu valor de brilho (escala de cinzentos) e cores que sejam muito escuras não necessitam de tantas distinções entre elas. Foram usados 18 contentores para o valor da cor (H), 4 para o valor da saturação (S) e 4 para o valor do brilho (B). De forma a reduzir o número de cores usadas, uma vez que muitas dificilmente são distinguíveis pelo olho humano, quando o brilho tem um valor muito baixo é contado como preto e se a saturação tiver um valor muito baixo podemos usar apenas o valor do brilho, ficando com 4 valores na escala de cinzentos: preto, cinzento escuro, cinzento claro e branco.

Ao oferecer os dois espaços de cor no VideoAnalyzer, o utilizador pode comparar os resultados de analisar o mesmo vídeo em espaços diferentes e escolher os melhores resultados em termos de percepção. Nas experiências conduzidas, o HSB ofereceu uma melhor percepção das cores para pesquisa, na criação de histogramas e para encontrar cores dominantes. Oferece uma melhor organização do espaço tornando mais fácil a

tarefa de agrupar as cores semelhantes quando se criam histogramas de cor. Estes resultados vão de encontro ao discutido em [Kerminen99, Saykol04].

As diferentes métricas de cor também podem ser comparadas, para ajudar a decidir a melhor forma de representação em cada situação.

Em vídeos onde algumas cores assumem relevância ao nível das frames, a contagem das cores dominantes por frame aparece naturalmente como uma boa métrica, uma vez que estas cores são normalmente perceptíveis enquanto se vê o vídeo e são bem aceites pelos utilizadores como dominantes. Em vídeos heterogéneos com muitas cores, a predominância das cores nas frames perde informação importante – uma cor pode aparecer muita vez mas nunca ser dominante numa frame. Neste caso, ao contar todos os píxeis em cada frame do vídeo temos maior precisão para obter as cores dominantes, com algumas reservas. Estas incluem tipicamente ter demasiadas cores dominantes, muito perto em termos percentuais umas das outras, para que alguma possa ser chamada de dominante. Para percebermos se um vídeo tem muitas cores dominantes é usada a métrica da percentagem de predominância das cores mais dominantes.

Na Figura 11, o primeiro gráfico da esquerda apresenta as cores dominantes no vídeo usando o método de contagem da cor de todos os píxeis de todas as frames do vídeo. Os resultados são muito semelhantes para algumas cores, frisando que estas serão as dominantes, mas sem que uma se distinga muito. Complementando esta informação, no segundo gráfico são apresentadas as cores dominantes contando-as apenas quando dominam numa frame individualmente. Aqui o vermelho acastanhado assume alguma preponderância ao dominar em cerca de 45% das frames. No terceiro gráfico podemos ver a predominância das cores mais dominantes nas frames onde são dominantes, o que costuma ser mais aproximado da nossa percepção de dominância.

Para permitir a selecção de boas representações para as cores dominantes, o VideoAnalyzer apresenta as diferentes métricas de uma maneira visual, deixando que os utilizadores percebam e possam escolher os resultados mais adequados a cada situação. Está também em estudo uma forma de sugestão automática das métricas a usar em cada situação, baseada nas propriedades do vídeo e na percepção dos utilizadores.

5. CONCLUSÕES E PERSPECTIVAS FUTURAS

Apresentámos o ColorsInMotion, uma aplicação que suporta a visualização e a exploração interactiva e criativa de vídeos, com ênfase nas propriedades de cor e movimento, cruzando acesso a informação, cultura e arte digital. Os utilizadores podem visualizar o espaço de vídeos através de diferentes vistas, para pesquisar, comparar e interagir com uma selecção de vídeos num contexto cultural, em áreas como a música e dança, de vários autores e países. Criámos vistas complementares baseadas no conteúdo dos vídeos, nas suas cores dominantes, no seu movimento, como os loops de vídeo, formas coloridas e slit scans, permitindo visualizar e

realçar algumas propriedades dos vídeos, simplificando a complexa informação que contém de uma forma perceptiva. Em termos de percepção, a escolha do espaço de cores e da estratégia de quantização foi um dos aspectos mais importantes na classificação e pesquisa. Os vídeos são pré-processados e indexados no sistema de acordo com várias dimensões, mas podem vir de várias fontes do mundo virtual ou real.

Este tipo de técnica está a tornar-se cada vez mais importante, à medida que a informação baseada em vídeo se vai tornando mais difundida. Pesquisas baseadas em anotações de utilizadores é normalmente suportada por repositórios como o YouTube, mas propriedades visuais como a cor e o movimento não são normalmente tidas em conta. Tendo os resultados do contexto correcto após pesquisa através destas anotações, as técnicas de visualização permitem que se apresente a informação de forma significativa, facilitando a percepção, comparação e procura através de navegação. No ColorsInMotion, a visualização pode por exemplo permitir: que um utilizador perceba e compare as diferentes cores e movimentos que existem no Fandango e nas danças Sevilhanas; ajudar a encontrar um vídeo de um cantor com predominância de tons azuis, juntamente com outros vídeos de cores semelhantes para incluir numa campanha publicitária; permitir que um video jockey faça uma selecção e composição de vídeos tendo em conta as suas propriedades visuais; ou permitir a criação interactiva de vistas com estéticas influenciadas pela cor e movimento como as que foram apresentadas na secção 3.4.

Esta aplicação pode ser usada num ambiente tradicional com um ecrã, teclado e rato, ou com um ecrã sensível ao toque, adequado em muitos contextos. Estão em desenvolvimento interfaces mais naturais, recurso a dispositivos, mais adequada a certos contextos de instalação e interacção ambiental. Juntamente com a captura de vídeos, cores e imagens do mundo real para efeitos de captura de informação e pesquisa, estão a ser afinadas interfaces baseadas no reconhecimento de cores e gestos para oferecer experiências mais ricas e imersivas de visualização e interacção baseadas em cor e movimento entre os espaços reais e virtuais.

Como perspectivas futuras pretendemos refinar o processamento e visualização de vídeos com base no resultado de avaliações mais sistemáticas, explorar novas propriedades do vídeo e vistas para diferentes contextos, bem como evoluir no sentido de tornar a visualização e interacção mais imersiva. A pesquisa semântica também pode evoluir para novos níveis no ColorsInMotion e, em cooperação com repositórios online, aceder a vídeos partilhados em cenários mais alargados, usando por exemplo serviços da Web 2.0 e metadados normalizados. Como exemplo, os mosaicos de vídeo poderiam ser construídos a partir de pesquisas baseadas em reconhecimento de conteúdo das imagens, bem como de metadados. Complementando as propriedades visuais com semânticas reduz-se o âmbito das pesquisas, levando a que os mosaicos tenham uma dimensão semântica mais significativa. Por exemplo, poderia ser dada uma imagem

da torre Eiffel, e os vídeos escolhidos para criar o mosaico seriam ou da própria torre Eiffel, de Paris ou de França, dependendo do âmbito e contexto. Para complementar a captura e influência do contexto do mundo real, as pesquisas poderiam ter em conta a localização do utilizador como modo adicional de filtrar a informação a aceder.

Outra perspectiva inclui estender a navegação interactiva para suportar edição de vídeo, usando para isso a experiência prévia com edição evolutiva de vídeo [Chambel07]. Edição de vídeo digital é uma tarefa que consome bastante tempo. As técnicas de pesquisa e visualização podem ser usadas para ajudar os utilizadores a escolher segmentos de vídeo a usar. Estas funcionalidades podem fazer com que a tarefa de edição seja feita mais rapidamente e de forma mais agradável.

6. AGRADECIMENTOS

O trabalho apresentado neste artigo foi parcialmente financiado pelo LaSIGE através do Programa de Financiamento Plurianual da FCT.

7. REFERÊNCIAS

- [Bestiario08] Bestiario, Videosphere, May 2008.
<<http://www.bestiario.org/research/videosphere/>>
- [Boreczky00] J. Boreczky, A. Girgensohn, G. Golovchinsky, and S. Uchihashi. An interactive comic book presentation for exploring video In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 01 - 06). CHI '00. ACM, New York, NY, 185-192, 2000.
- [Chambel02] T. Chambel, and N. Guimarães, Context Perception in Video-Based Hypermedia Spaces. In Proceedings of ACM Hypertext'02, College Park, Maryland, USA, 11-15 June, 2002.
- [Chambel07] T. Chambel, L. Correia, J. Manzolli, G.D. Miguel, N.A. Henriques, and N. Correia. Creating Video Art with Evolutionary Algorithms. Special Issue on Technology and Digital Art, *Computer Graphics* 31, 6, pp.837-847, Dec 2007.
- [Daniel03] G. Daniel, and M. Chen. Video Visualization In Proceedings of the 14th IEEE Visualization 2003 (Vis'03) (October 22 - 24). IEEE Visualization. IEEE Computer Society, Washington, DC, 54, 2003.
- [Fels99] S. Fels, and K. Mase, Interactive video cubism. In Proceedings of NPIVM '99, Workshop on New Paradigms in Information Visualization and Manipulation (Kansas City, Missouri, USA, Nov 2-6), ACM, New York, NY, 78-82, 1999.
- [Few07] S. Few. Data Visualization: Past, Present, and Future. IBM Cognos Innovation Center, 2007.
- [Finkelstein98] A. Finkelstein, and M. Range. Image Mosaics. In Proceedings of the 7th international Conference on Electronic Publishing. R. D. Hersch, J. André, and H. Brown, Eds. Lecture Notes In Computer Science, vol. 1375. Springer-Verlag, London, 11-22, 1998.
- [Girgensohn00] A. Girgensohn, J. Boreczky, P. Chiu, J. Doherty, J. Foote, G. Golovchinsky, S. Uchihashi, and L. Wilcox. A semi-automatic approach to home video editing. In Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology (San Diego, CA, USA, Nov 6-8). UIST '00. ACM, New York, NY, 81-89, 2000.
- [Idée08] Idée's Multicolr Search Lab, 2008.
<<http://labs.ideeinc.com/multicolr/>>
- [Irani95] M. Irani, P. Anandan, and S. Hsu. Mosaic based representations of video sequences and their applications. In Proceedings of the Fifth international Conference on Computer Vision (June 20 - 23, 1995). ICCV. IEEE Computer Society, Washington, DC, 605, 1995.
- [Kerminen99] P. Kerminen, and M. Gabbouj. Image Retrieval Based on Color Matching, Proceedings of FINSIG'99, Oulu, Finland, 89-93, 1999.
- [Klein02] A.W. Klein, T. Grant, A. Finkelstein, and M.F. Cohen. Video mosaics. In Proceedings of the 2nd international Symposium on Non-Photorealistic Animation and Rendering (Annecy, France, June 03 - 05, 2002). NPAR '02. ACM, New York, NY, 21-28, 2002.
- [Kneller65] G.F. Kneller. Art and Science of Creativity, International Thomson Publishing, 1965.
- [Levin08] G. Levin and Collaborators Catalogues and Lists, An Informal Catalogue of Slit-Scan Video Artworks and Research, 2005-2008.
<http://www.flong.com/texts/lists/slit_scan/>
- [Makela06] M. Makela. LIVE CINEMA: Language and Elements. MA in New Media, Media Lab, Helsinki University of Art and Design, 2006.
- [Manovich] L. Manovich, and A. Kratky. Soft Cinema: Ambient Narrative, <<http://softcinema.net>>
- [Martin04] S. Martin. Video Mosaics, Berkeley, Final Project for CS283 Graduate Graphics with Prof James O'Brien, 2004.
<<http://stevezero.com/eecs/mosaic/index.htm>>
- [Nunes07] M. Nunes, S. Greenberg, S. Carpendale, and C. Gutwin. What did I miss? Visualizing the past through video traces. Proc ECSCW'07 European Conf on Computer Supported Cooperative Work, Springer-Verlag, 2007.
- [Paul03] C. Paul. Digital Art. London: Thames & Hudson, 2003.
- [Pepperell02] R. Pepperell. Computer aided creativity: practical experience and theoretical concerns In Proceedings of the 4th Conference on Creativity & Cognition (Loughborough, UK, October 13 - 16). C&C '02. ACM, New York, NY, 50-56, 2002.
- [Perez08] S. Perez. The Best Tools for Visualization, 2008.<http://www.readwriteweb.com/archives/the_best_tools_for_visualization.php>

[Rocha08] T. Rocha, and T. Chambel. VideoSpace: A 3D Video Experience In Proceedings of Artech 2008, the 4th International Conference on Digital Arts, pp. 305-310, Porto, Portugal, November 7-8, 2008.

[Saykol04] E. Saykol, U. Güdükbay, and O. Ulusoy. Integrated Querying of Images by Color, Shape, and Texture Content of Salient Objects. In Lecture Notes in Computer Science (LNCS), Vol. 3261, Advances

in Information Sciences (ADVIS'2004), Edited by Tatyana Yakhno, pp. 363-371, Springer-Verlag, Izmir, Turkey, 2004.

[Sims91] K. Sims. Artificial evolution for computer graphics In Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '91. ACM, New York, NY, 319-328, 1991.