Content Based Video Query

ALEXANDRE FERREIRA[1], PAULO MATTOS[1], RODRIGO TOLEDO[1], JONAS GOMES[2], AND LUIZ VELHO[2]

[1]TeCGraf - Grupo de Tecnologia em Computao Grfica, Departamento de Informtica, PUC-Rio
{alexgf, pmattos, rtoledo}@tecgraf.puc-rio.br
[2]VISGRAF Laboratory, IMPA, Rio de Janeiro, Brazil
{Jonas, lvelho}@visgraf.impa.br

1. The Problem

The use of image, audio and video databases is assuming a growing importance in many applications such as GIS, distance learning, television production, medical images, Information technology and so forth. An important and challenging problem is the query of these databases by content (QBC). This poster describes a project to develop a QBC system for video databases. The project extends to video the technique introduced in "Fast Multiresolution Image Querying" by Salesin et al. for a QBC system for an image database using wavelets.

2. Image Query

The key to the algorithm is the establishment of an effective and efficient metric capable of computing the distance between a query image Q and a potential target image T. Wavelet decomposition proved to be a good foundation for this metric, for several reasons, such as: few coefficients provides a good approximation of the original image retaining information from existing edges (see figure 1); presents relative invariance to resolution changes; it is fast to compute, running in linear time in size of the image; spatial localization of the frequencies. The chosen metric use the YIQ color space and the Haar wavelet.

For each image we compute its Haar wavelet transform, truncating and quantizing its coeffecients. Those remaining coeffecients represent the image signature.

The metric for each color channel can be represented as:

$$||Q,T|| = w_{0,0}|Q[0,0]-T[0,0]| + \sum_{i,j} w_{i,j}|Q'[i,j]-T'[i,j]|$$

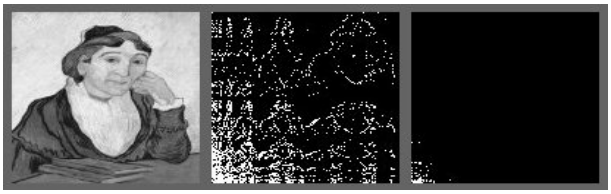

Figure 1 - An original image; its signature with 2000 and 60 coefficients

Where $Q[0,0]$ and $T[0,0]$ correspond to the overall average intensity of that color channel; $Q'[i,j]$ and $T'[i.j]$ represent the $[i,j]$-th truncated (the mth greatest), quantized (-1, 0, 1) wavelet coefficients (terms) of Q and T; and $w_{i,j}$ the weight of the $[i,j]$-th coefficient.

The application implements both the pre-processing and query interface of the algorithm. The pre-process phase consists in creating an image database containing the signatures of each image. The program offers two query options: the user inputs a pre-existing image or the user draws a sketch of the intended image on the canvas (see figure 2).

Figure 2

3. Video Query

The image query algorithm described above can be extended to volumetric graphical objects. The extension is straightforward:

- The volumetric object is given by a matrix representation (voxels);
- We extend the Haar transform to volumetric objects using tensor product;
- The computation of the volumetric signature is also a simple extension of the 2D solution.

The video query consists in considering a video sequence as a volumetric object (see Figure 3). Two possibilities for que input video query are possible: Use a video shot of the desired sequence (a time normalization should be considered). Use a paint system interface where user strokes, represents the motion path of an essential motion in the required video sequence.

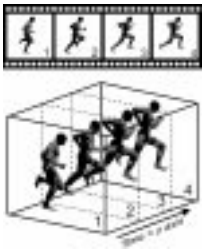

Figure 3