

基于内容的电影动画素材检索

李 伟, 王树梅, 王 玲

(南京理工大学计算机学院, 南京 210094)

摘 要: 传统的电影动画素材检索, 通常采用基于索引的方法。素材库中的动画素材数量巨大, 采用这种方法会造成人力和时间的浪费, 达不到满意的效果。MPEG-7 标准是多媒体内容描述的接口, 该文将 MPEG-7 标准运用到电影动画素材的检索中, 缩短了制作周期, 提高了工作效率。

关键词: MPEG-7; 颜色描述符; 纹理描述符; 动画素材检索

Content-based Film Animation Material Search

LI Wei, WANG Shumei, WANG Ling

(School of Computer, Nanjing University of Science and Technology, Nanjing 210094)

【Abstract】 The traditional way to search animation material is usually based on index, which may be time-consuming and can not achieve the satisfying outcome, for there are numerous materials in database. Using the methods supplied by the MPEG-7 standard, the interface of describing multimedia content, to search animation materials, reduces the manufacture cycle greatly and enhances the working efficiency.

【Key words】 MPEG-7; Color descriptor; Texture descriptor; Animation material search

在制作电影的过程中, 快速检索所需要的素材是非常必要的。传统的基于索引的动画素材检索方法无论从效率还是准确度上都不能满足电影工作者的要求, 人们迫切需要一种快速、准确的方法查找到素材, 而基于内容的多媒体检索方法可以解决这类问题。

2001年7月, 国际标准——MPEG-7 提供了一整套多媒体内容描述工具, 进一步发展成基于内容的描述和检索规范, 它以固定的格式储存图像或声音的内容特征。制作动画人员可以预先以 MPEG-7 标准格式储存动画素材, 并提供给检索系统使用, 以达到快速检索的效果。

由于已有的动画素材通常都是一段镜头, 该段镜头是由一系列视觉上变化不大的图像帧组成, 因此, 可以用检索关键帧的方法来检索动画素材。在 MPEG-7 标准中通常用颜色和纹理描述符来提取此类静态图像的特征。在制作动画的过程中, 植物、白云等这类关键帧的颜色和纹理特征均很明显, 用单一描述符来描述它们的内容, 实验得到的检索效果不是很理想。本文的主要工作是结合这两类描述符的优点, 进行实验。实验结果表明: 查准率和查全率与单一描述符相比均有提高。

1 颜色特征

颜色作为图像最基本的而且是最具表现力的视觉特征, 具有定义比较明确、抽取相对容易的特点。基于颜色的图像检索得到了广泛的重视, MPEG-7 颜色描述符见图 1^[1]。

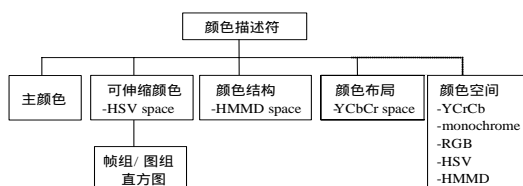


图1 MPEG-7 颜色描述符结构

由图1可以看出, 多个颜色“描述子”组成 MPEG-7 中的颜色“描述符”, 本文主要考虑其中主颜色“描述子”的应用。

1.1 主颜色描述子

在制作动画电影时通常着色不会很多, 一幅图用几种或十几种颜色信息, 而且根据习惯, 最重要的信息往往出现在图的中央。因为主颜色描述子主要用于描述图像中显著颜色的分布信息, 其目的是为兴趣区域提供一种有效、紧凑和直观的颜色表示, 所以其最适合用于表示局部(某个对象或图像区域)的显著颜色特征。

1.2 主颜色描述子的检索算法

1.2.1 提取方法

(1) 将图像从RGB空间转换到HSV空间^[2]。

(2) 考虑到计算代价, 需要将HSV空间进行量化。量化的方法有很多, 一般采用比较成熟的72色或36色非均匀量化算法^[6,7]。为了能进一步地提高检索速度和准确率, 本文采用比较先进的22色非均匀量化算法。

(3) 将一幅图分成4×4的子块, 如图2所示, 因为用户通常对中间区域的信息感兴趣, 所以中间所占的权重稍多些^[8]。实验表明权重为A:B:C=6:4:1时, 效果最佳。

C	B	B	C
B	A	A	B
B	A	A	B
C	B	B	C

图2 图像分割区域

作者简介: 李 伟(1981-), 男, 硕士生, 主研方向: 模式识别, 人工智能; 王树梅, 教授; 王 玲, 副教授

收稿日期: 2006-08-02 **E-mail:** liwei81730@163.com

(4)计算出每种颜色所占整幅图的百分比。用符号 FC 代表描述符,它由颜色和对应的百分比构成,即

$$FC=\{[Ci, Pi], i=0, \dots, N, Pi \in [0,1]\}$$

其中, N 代表颜色的数目; Ci 表示颜色; Pi 表示颜色所占的百分比。当 $Pi>5\%$ 或 $Pi>60\%$ 时,则把该颜色作为主颜色。

1.2.2 相似度的测量

2 个主颜色“描述子”为

$$FC_1=\{[Ci, Pi], i=0, \dots, M, Pi \in [0,1]\}$$

$$FC_2=\{[Cj, Pj], j=0, \dots, N, Pj \in [0,1]\}$$

2 个描述子之间的距离为

$$D_C(FC_1, FC_2) = \sum_{i=1}^M \sum_{k=1}^M a_{i,k} P_i P_k + \sum_{j=1}^N \sum_{l=1}^N a_{j,l} P_j P_l - \sum_{i=1}^M \sum_{j=1}^N 2a_{i,j} P_i P_j$$

其中,系数 $a_{i,j}$ 表示两种颜色的相似性,即

$$a_{i,j} = 1 - d_{i,j} / d_{\max}$$

其中, $d_{i,j}$ 表示两种颜色之间的欧式距离; d_{\max} 则是两种颜色之间的最大距离。通过计算两种完全相反的颜色可得出 $d_{\max} = \sqrt{5}$ 。

1.2.3 实验结果

从动画素材库中的景物类素材库中随机抽取包括晚霞、水波、鲜花、白云这 4 类景色,每类景色抽取 30 幅关键帧(也就是代表 30 个镜头),共计 120 幅关键帧,作为实验的基础数据库。实验时,每类题材景色的关键帧都将作为待检索图像,用主颜色描述符在数据库中进行检索。图 3 是查询的反馈结果,排在前面的是相似度高的图像。



图 3 主颜色描述子检索的结果

此次实验的评价标准是多次检索的查准率和查全率,即:

查准率 $P = \text{检索出的相关图像数} / \text{被检索出的图像总数} = 72.4\%$

查全率 $R = \text{检索出的相关图像数} / \text{所有图像总数} = 31.6\%$

2 纹理特征

纹理是静态图像的底层特征之一,它表示一幅图像的大量相似的图样(pattern),在动画素材中也有大量纹理特征明显的镜头,例如水波、鲜花、家具表面的木纹等图样元素决定了纹理图像的外观,可以利用图样元素的排列方式定义纹理图像的方向、大小和规则度,并以此来检索相似的纹理图像或浏览具有相同方向、大小和规则度的纹理图像。

在 MPEG-7 中定义了 3 种描述纹理图像特征的方法,即同构型纹理图像描述子、纹理图像浏览描述子和边界直方图描述子。本文主要用到的是同构型纹理图像描述子。

2.1 同构型纹理图像描述子

同构型纹理图像描述子(homogeneous texture descriptor, HTD)由 62 个数字组成^[3],要根据图像纹理的方向和大小提取底层特征。在 62 个数字中,开头的 2 个分别代表图像空间

域中的平均值和标准差,由符号 f_{CD} 、 f_{SD} 表示,其余的 60 个数字则代表不同频率的频道上能量的平均值和标准差,由 Gabor 滤波器来增强这些频道上的能量,图 4 显示了频率域的分布情况^[3]。

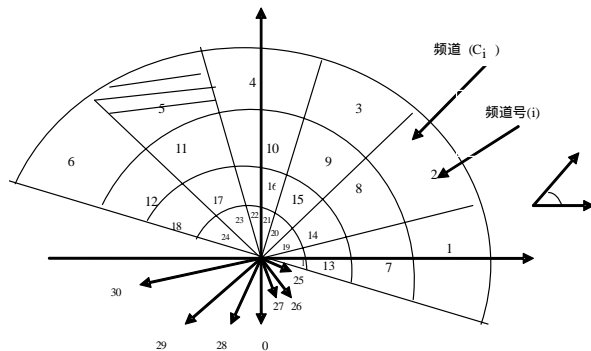


图 4 频率域的分布

将频率域按照方向和半径分割成 30 个频道,频道在角度上是以 30° 等量分割为 6 个方向,而在半径上则是以 $1/2$ 为底,等比分割为 5 个半径,然后计算在每个频道中的平均能量和其方差值,描述子可以表示为

$$HTD = \{f_{CD}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}\}$$

其中, e_i 和 d_i 分别表示第 i 个频道中的能量平均值和能量标准差。

2.2 同构型纹理图像描述子的检索算法

f_{CD} 和 f_{SD} 的提取方法很简单,直接利用像素值计算就可以了。本文主要讨论如何提取 i^{th} 频道的能量平均值和标准差。

2.2.1 提取方法

(1)将图像从 RGB 空间转换到 HSV 空间^[2]。由于亮度信息基本包含了图像的纹理信息,因此,本文只用该分量作为该图像的主要纹理信息。

(2)将输入的图像经过一维傅里叶变换的方法变换到频域中,用极坐标的方式 $F(\cdot, \cdot)$ 表示。

(3)用 Gabor 滤波器处理变换后的结果,找出图像的方向特性。

Gabor 滤波器可以加强频率域中某个方向和大小上的亮度,如果输入的图像中的纹理方向和大小同 Gabor 滤波器所使用的方向吻合,则可以得到最大的亮度^[4],利用这个特点,可以计算图像中的纹理所含有的方向和大小。

经过 Gabor 滤波器处理后的结果为

$$H_i(\cdot, \cdot) = G_s(r(\cdot, \cdot)) F(\cdot, \cdot), i = 6 \times s \times r + 1$$

其中, i 表示频道的索引; $G_s, r(\cdot, \cdot)$ 是第 s 个半径和第 r 个方向分割上的 Gabor 函数,即

$$G_s, r(\cdot, \cdot) = \exp\left[-\frac{(\omega - \omega_s)^2}{2\sigma_s^2}\right] \cdot \exp\left[-\frac{(\theta - \theta_r)^2}{2\sigma_r^2}\right]$$

其中, $\sigma_s = \frac{B_s}{2\sqrt{2\ln 2}}$, $\sigma_r = \frac{B_r}{2\sqrt{2\ln 2}}$ 。

B_s 和 B_r 是频率域中分割的半径分量和角度分量的间距。

表 1、表 2 给出了 Gabor 函数所用到的参数^[5]。

表 1 半径分量所用的参数

半径索引(s)	0	1	2	3	4
中心频率(ω_s)	3/4	3/8	3/16	3/32	3/64
间距(B_s)	1/2	1/4	1/8	1/16	1/32
σ_s	$\frac{1}{4\sqrt{2\ln 2}}$	$\frac{1}{8\sqrt{2\ln 2}}$	$\frac{1}{16\sqrt{2\ln 2}}$	$\frac{1}{32\sqrt{2\ln 2}}$	$\frac{1}{64\sqrt{2\ln 2}}$

表 2 角度分量所用的参数

角度索引(s)	0	1	2	3	4	5
中心频率(r)	0°	30°	60°	90°	120°	150°
间距(Br)	30°	30°	30°	30°	30°	30°
σ_r	$\frac{30^\circ}{2\sqrt{2}\ln 2}$	$\frac{30^\circ}{2\sqrt{2}\ln 2}$	$\frac{30^\circ}{2\sqrt{2}\ln 2}$	$\frac{30^\circ}{2\sqrt{2}\ln 2}$	$\frac{30^\circ}{2\sqrt{2}\ln 2}$	$\frac{30^\circ}{2\sqrt{2}\ln 2}$

(4)计算平均能量 e_i 和能量标准差 d_i , 即

$$e_i = \log[1 + p_i], p_i = \sum_{\omega} \sum_{\theta} H_i(\omega, \theta)$$

$$d_i = \log[1 + q_i], q_i = \sqrt{\sum_{\omega} \sum_{\theta} [H_i^2(\omega, \theta) - p_i]^2}$$

根据上述公式可以计算出 30 个频道的平均能量和其能量标准差, 连同 f_{CD} 和 f_{SD} 一起记录在 $HTD = \{f_{CD}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}\}$ 中, 用 HTD 就可以描述一幅图像的纹理。

2.2.2 相似度的测量

利用两幅图像纹理的差异距离为

$$D_T(HTD_1, HTD_2) = \sum_k \left| \frac{HTD_1(k) - HTD_2(k)}{\alpha(k)} \right|$$

测量它们之间的相似度, 差异距离越大相似度越低, 差异距离越小相似度越高。其中, k 为整数, 范围为 1~62; $HTD(k)$ 表示某图像的 k^{th} 个 HTD 的值; $\alpha(k)$ 是正规化参数, 可以使用查询数据库中所有图像的第 k 个 HTD 值计算出的标准差。这里为了计算简单, 本文规定 $\alpha(k) = 1$ 。

2.2.3 实验结果

在第 1 个实验使用的数据库的基础上, 用同构型纹理图像描述子在数据库中进行检索, 图 5 显示了反馈结果。

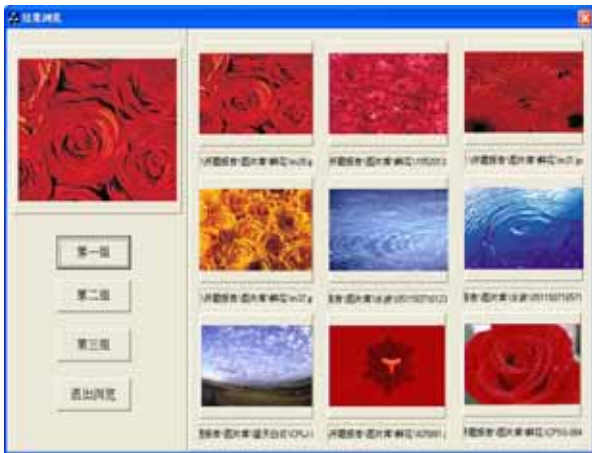


图 5 纹理描述子检索的结果

用多次实验所得的查准率和查全率作为评价标准, 即:

查准率 $P = \text{检索出的相关图像数} / \text{被检索出的图像总数} = 77.3\%$

查全率 $R = \text{检索出的相关图像} / \text{所有的相关的图像} = 33.2\%$

3 综合检索

在电影动画素材中, 有些镜头的关键帧的颜色和纹理特征都很明显, 此时如果能将主颜色描述子和同构型纹理图像描述子的优点结合在一起, 成为判断相似度的标准, 将进一步提高检索的准确率。

3.1 检索方法

本文采用人工检索和自动检索相结合的方法, 由检索者根据待检索关键帧的内容为主颜色描述子、同构型纹理图像描述子, 设定相应的权重, 权重的大小将影响到检索的结果。考虑到动画的制作是一项艺术的创作, 制作者的主观感觉也占有重要的地位, 由制作者自己决定待检索关键帧是颜色信

息重要还是纹理信息重要。

3.2 相似度的测量

假设使用主颜色描述子得到的两幅图像帧的差异距离为 D_1 , 使用同构型纹理图像描述子得到的差异距离是 D_2 , 为了使 D_1 和 D_2 具有可比性, 本文进行了归一化处理, 即

$$D_C(FC_1, FC_2) = \frac{MAX_1 - D_1}{MAX_1}$$

其中, MAX_1 是主颜色描述子差异距离的最大值。

$$D_T(HTD_1, HTD_2) = \frac{MAX_2 - D_2}{MAX_2}$$

其中, MAX_2 是同构型纹理图像描述子差异距离的最大值。

$D_C(FC_1, FC_2)$ 和 $D_T(HTD_1, HTD_2)$ 的范围为 0~1。

综合主颜色描述子和同构型纹理图像描述子的相似度测量方法, 可以由下式表示:

$$D = \alpha D_C(FC_1, FC_2) + \beta D_T(HTD_1, HTD_2), \alpha + \beta = 1$$

其中, α 和 β 就是为它们设置的权重; D 越大, 相似度越低; D 越小, 相似度越高。

3.3 实验结果

与上述 2 个实验的条件相同, 由检索者根据自己的感觉为颜色和纹理描述符设置相应的权重。图 6 是实验的效果截图。



图 6 综合颜色和纹理描述符的检索

多次实验所得的查准率和查全率为

查准率 $P = \text{检索出的相关图像数} / \text{被检索出的图像总数} = 84.7\%$

查全率 $R = \text{检索出的相关图像} / \text{所有的相关的图像} = 39.3\%$

从实验结果可以看出, 将颜色和纹理描述符的优点结合在一起进行检索, 要比单纯用颜色描述符和纹理描述符进行检索的查准率和查全率都要高。实验证明, 综合颜色和纹理的检索准确率确实有所提高。

4 结论

虽然综合颜色和纹理特征对动画关键帧进行检索取得了良好的实验结果, 但是这种方法也是有局限性的。实验表明, 如果是纯景物类的关键帧, 用颜色特征进行检索, 效果会更好些。像家具表面木纹这样的纯纹理关键帧, 用纹理特征检索, 效果较好。这就要求检索者在检索过程中要根据实际情况, 灵活地设定颜色和纹理描述符的权重, 以达到较好的检索结果。另外, 对人物形象出现比较多的关键帧, 用颜色和纹理特征进行描述, 检索的效果并不理想。如何进一步扩大检索关键帧内容的范围, 是下一步研究的方向。

(下转第 230 页)