

基于分形纹理特征和小波变换的网状纹理检测方法

曹媛媛¹ 杨波¹ 徐光祐¹

摘要 网状纹理广泛存在于日常生活中, 其独特的网格结构以及极易发生形变的特点使网状纹理的描述和检测成为一件困难的工作. 本文以球网检测为例, 提出了基于分形纹理特征和小波变换的网状纹理检测方法. 先对足球视频图像进行小波多分辨率分解, 计算不同尺度图像的分形纹理特征向量, 基于因果关系对特征进行选择组成特征向量. 然后使用支持向量机检测图像中的网状纹理. 实验表明此方法有较强的鲁棒性, 能够在剧烈形变、复杂背景和基元大小差异明显的条件下成功检测出球网.

关键词 网状纹理, 球网检测, 分形纹理特征, 多分辨率分析, 特征选择
中图分类号 TP391

Netlike Texture Detection Method Using Fractal Texture Features and Wavelet Transform

CAO Yuan-Yuan¹ YANG Bo¹ XU Guang-You¹

Abstract Netlike texture is very common in our daily life. However, it is very difficult to be described and detected because of its distinct characteristics. This paper chooses nets in sports game as research object and presents a netlike texture detection method of using fractal texture features and wavelet transform. First, multi-scale wavelet decomposition is employed on video frames; then fractal features of multi-resolution images are obtained. A group of fractal features are chosen as feature vector according to causality theory and an SVM classifier is used to recognize the net in images. The experiments show that this method is robust and effective under the condition of intense distortion and complicated background. This method can be extended to other netlike texture detections.

Key words Netlike texture, net detection, fractal texture feature, multi-resolution analysis, feature selection

1 引言

网状纹理是现实生活中普遍存在的一种纹理, 例如各种比赛中的球网、渔网、纱窗以及布料等, 网状纹理的检测和识别将为图像识别和理解提供有力的线索, 因此它具有普遍而现实的意义. 目前对网状纹理的检测和识别没有受到足够的重视, 其原因是网状纹理的出现大多伴随有形变的发生, 而且常常与背景相混合, 使得网状纹理的检测极易受到背景的干扰, 难以描述. 本文中对网状纹理的检测用球网做为研究对象. 因为球网作为一个具有丰富语义信息的最直接的上下文线索, 在基于内容的体育视频事件检索中有重要的应用, 然而目前对球网的研究却很少. 球网具有典型的网状纹理的特征, 形态特征差异较大, 难以找到统一的描述. 这种差异有三方面: 1) 球网常因球员或球的撞击而发生不可预知的

形变; 2) 不同比赛中的球网网孔的形状不同, 有四边形、六边形等; 3) 光照条件与摄像机成像条件不同造成的球网颜色不稳定.

本文针对网状纹理具有强烈的自相似性, 不同尺度的特性, 提出了把分形特征与多分辨率小波变换相结合的网状纹理检测方法. 在不同分辨率的图像上抽取分形纹理特征, 并进一步根据因果关系理论, 从一组分形纹理特征中选取对球网具有较强描述能力的特征组成特征向量. 最后采用基于支持向量机 (Support vector machine, SVM) 的方法将图像分为有球网和无球网两类. 在不同球类体育视频上的实验结果表明该方法能够快速准确地检测出体育视频中的球网, 证明了分形纹理特征对网状纹理具有很强的描述能力. 此方法也可以推广用于其它有形变或复杂背景的网状纹理的检测. 本文第 2 节介绍分形纹理特征估计方法; 用因果关系来选择合适的分形纹理特征在第 3 节叙述; 第 4 节介绍对图像进行小波多分辨率分解.

2 分形纹理特征估计方法简介

2.1 分形维数的概念

1975 年, Mandelbrot 在其《自然界中的分形几何》中引入了分形 (fractal) 这一概念. Fisher^[1] 给

收稿日期 2005-11-15 收修改稿日期 2006-7-1
Received November 15, 2005; in revised form July 1, 2006
国家自然科学基金 (60273005) 资助
Supported by National Natural Science Foundation of P. R. China (60273005)
1. 清华大学计算机科学与技术系普适计算教育部重点实验室 北京 100084
1. Key Laboratory of Pervasive Computing, Ministry of Education, Department of Computer Science and Technology, Tsinghua University, Beijing 100084
DOI: 10.1360/aas-007-0688

出了分形的一个描述性定义: 如果一个集合 F 被视为分形, 那么认为它应具有以下 (部分) 特点: 1) F 在任意尺度都存在细节; 2) F 具有完全的 (或部分的、概率的) 自相似性; 3) F 的分形维数大于其拓扑维数; 4) F 有一个简单的算法性描述 (迭代函数系 IFS). 对于分形维数, Mandelbrot^[2] 定义为: 唯一的能够产生一集合测度特性的恒定估测的分形幂, 被称为这个集合的分形维数. 此外, 基于不同的测度和分形构造方法, 以及应用领域的差异, 人们还定义了其他一些分形维数, 如自相似维数, 盒维数, Housdorff 维数等等. 实际应用中, 由于盒维数易于估计, 所以成为常用的分形特征.

2.2 盒维数估计方法

在分形维数的估计上, 本文分别选用 Keller 和 Chaudhuri 两人提出的分形盒维数估测方法^[3,4]. Keller 方法首先将图像 A 视为一个三维空域中的一个表面 $(x, y, f(x, y))$, 其中 $f(x, y)$ 为图像在 (x, y) 位置处的灰度值. 令 $P(m, L)$ 为以图像表面中任一点为中心的大小为 L 的盒子中, 点的数目为 m 的概率. 令 N 为盒子中最多可容纳的点的数目, M 为图像中点的数目. 所以含有 m 个点的盒子的数目为 (M/m) . 这样, 用来覆盖整个图像所需的盒子数的期望值为

$$\langle N(L) \rangle = \sum_{m=1}^N (M/m)P(m, L) = M \sum_{m=1}^N (1/m)P(m, L) \quad (1)$$

这样, 如果令

$$N(L) = \sum_{m=1}^N (1/m)P(m, L) \quad (2)$$

可得

$$N(L) \propto L^{-D} \quad (3)$$

其中 D 为定义的分形维数, 且有 $D \in [2.0, 3.0]$. 在实际计算中, 通过 $P(m, L)$ 的计算, 可以求得一系列点对 $(\log(L), \log(N(L)))$. 由式 (3) 可知, 这些点呈线性关系. 通过最小二乘拟合, 得出线性模型的方程, 其中直线的斜率即为待估测的分形维数.

Chaudhuri 指出分形维数的定义由下面的公式给出

$$1 = N_r r^D \quad \text{or} \quad D = \frac{\log(N_r)}{\log(1/r)} \quad (4)$$

公式中的 N_r 由以下方法确定: 假设一个 $M \times M$ 像素的图像被缩小至 $S \times S$ 大小, 且满足 $M/2 \geq s > 1, s \in I, r$ 被定义为 s/M . 然后将图像视为三维空域中的一个表面 $(x, y, f(x, y))$,

其中 $f(x, y)$ 为图像在 (x, y) 位置处的灰度值. 将 (x, y) 平面分成 $s \times s$ 的网格, 每个 $s \times s$ 的网格构成了一个柱体的底面, 这个柱体可以看作是由一系列 $s \times s \times s$ 的盒子组成. 将此 $s \times s$ 网格中所有像素的最小灰度值落在的盒子序号定为 k , 将最大灰度值落入的盒子的序号定为 l , 那么这个网格对 N_r 的贡献就可以记作 $n_r = l - k + 1$. 计算了所有网格的 n_r 值, 我们就可以得到 $N_r = \sum_{i,j} n_r(i, j)$, 然后根据分形维数的定义 (式 (4)), 在实际计算中, 通过 $\log(N_r)$ 和 $\log(1/r)$ 的计算, 可以得到一系列点对 $(\log(N_r), \log(1/r))$. 通过最小二乘拟合, 得出线性模型的方程, 其中直线的斜率即为待估测的分形维数. 据此可以估计出该图像的分形维数.

3 基于因果关系的特征选择

3.1 初始 14 个分形纹理特征的计算

为了提高分形纹理特征对纹理的描述和区分能力, 我们构造一组由原图像经过不同方向的低通和高通滤波的变换图像如下: 1) 水平方向平滑图像; 2) 垂直方向平滑图像; 3) 任意方向平滑图像; 4) 水平方向边缘图像; 5) 垂直方向边缘图像; 6) 任意方向边缘图像. 我们分别采用 Keller 和 Chaudhuri 两种不同的分形维数估计方法去估计以上 6 种变换图像和原图像的分形维数得到 14 个的分形纹理特征. 这些分形纹理特征不仅可以给出纹理基元的方向性的描述, 而且给出纹理基元稀疏性的描述. 有助于实现对网状纹理的充分描述.

但是在不同的检测对象和工作条件中, 这些分形纹理特征对检测结果的贡献是不一样的, 关系不大的特征的计算误差也可能造成检测错误. 而且, 对于每一种纹理特征的计算, 系统都有计算的代价, 特征向量维数过高也会使分类问题变得复杂, 所以我们需要一套方法来评估以上几种分形纹理特征和球网的相关程度, 进而选择那些真正对检测球网有重要价值的纹理特征.

3.2 互信息理论

我们可以用纹理特征提供的信息量来度量该纹理特征与球网之间的因果关系和相关程度^[5]. 按照信息论的观点, 这可以通过互信息 (Mutual information)^[6] 来计算

$$I(X, C_i) = \sum_X \sum_{C_i} P(X, C_i) \ln \frac{P(X, C_i)}{P(X)P(C_i)} \quad (5)$$

由上式可以看出, 如果 C_i 和 X 没有因果关系, 也就是说它们是独立的, 那么有 $P(X, C_i) = P(X)P(C_i)$, 从而 $I(X, C_i) = 0$. 一般地, C_i 和 X 的因果关系越强, 关联性越大, 它们的互信息

$I(X, C_i)$ 就越大.

因此, 某种分形纹理特征与球网间的互信息值的大小就表征了该纹理区分球网图像与非球网图像的能力的强弱. 通过因果关系的计算, 我们可以有效地选择那些与球网真正有关联的纹理特征, 这些纹理特征即是对球网描述能力最强的特征.

3.3 基于互信息的纹理特征选择

我们需要利用正负样本来估计互信息的值, 从而达到选择特征的目的. 令式 (1) 中 X 表示样本图像类别 (本文为正负两类), C_i 表示第 i 个特征 ($i = 1, \dots, 14$), $P(X, C_i)$, $P(X)$, $P(C_i)$ 可以通过样本在各维特征上的分布统计得到.

1) 样本选择和特征计算.

假定足球视频尺寸为 $M \times N$, 手工分割出 $M/2 \times N/2$ 大小的球网图像块作为正样本, 正样本需具有足够多的基元并且基元尺寸大致相同. 取观众、草地或者球员的图像块作为负样本. 正负样本集合各包含 100 个样本. 几个样本的例子见图 1, 分别计算 200 个样本的 14 个分形特征.

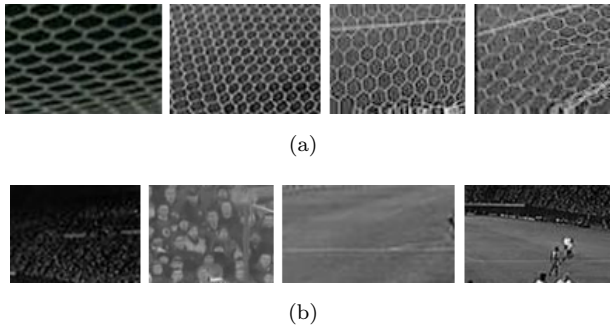


图 1 正样本与负样本举例 ((a) 正样本; (b) 负样本)

Fig. 1 Examples of positive and negative samples ((a) Positive samples; (b) Negative samples)

2) 分形纹理特征的选择.

按照式 (1) 计算每一维特征与样本类别之间的互信息值, 这个互信息值即表征了特征与类别间的相关性, 即该特征可以对正负样本正确分类的能力. 按互信息值由大到小的部分特征见表 1.

表 1 按相关性由大到小排列的前 5 个分形特征

Table 1 The first 5 fractal features which are of highest mutual information

	分行纹理特征	互信息值
1	Chaudhuri 任意方向平滑图像的 FD	0.419403
2	Chaudhuri 水平边缘图像的 FD	0.339009
3	Chaudhuri 原图像的 FD	0.297006
4	Chaudhuri 水平平滑图像的 FD	0.256902
5	Chaudhuri 垂直方向平滑图像的 FD	0.251284

从表 1 可以看出, 在检测球网纹理的应用中, 使

用 Chaudhuri 估计方法具有更好的分类效果. 出于减少计算量的考虑, 试验中只取前 5 个特征组成特征向量, 只使用 Chaudhuri 估计方法估计分形维数.

需要说明的是, 这里的计算是在假设各个特征关于球网条件独立的情况下进行的, 所以我们才可以逐一检查各个线索与事件的因果关系. 如果特征之间关于球网不是条件独立的, 那么仅仅通过这种方式来评估特征可能是不够的.

4 小波多分辨率分析 (MRA)

由于镜头拍摄尺度不同, 视频中球网网孔的大小也有很大的差异, 如图 2(a) 和图 2(b) 所示. 为了兼顾各个尺度的球网, 我们采用了小波多分辨率分析的方法, 将具有较大基元的网状纹理缩小到不同尺度上进行分形纹理估计.

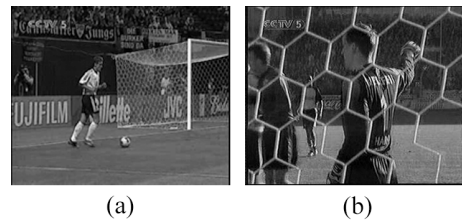


图 2 (a) 中镜头中的球网 (b) 近镜头中的球网

Fig. 2 (a) Net in zoom-in shot; (b) Net in close-up shot

Mallat^[7] 首先引入了多尺度分析的理论, 认为图像小波变换过程可以视为多通道多分辨率特征的提取过程, 小波分解的低频部分是图像在不同分辨率下的近似逼近, 描述了图像在不同尺度下的全局概貌特征; 小波分解的高频部分描述了图像在不同方向上灰度变化特征. 同时, 图像小波多尺度分解描述的是一种变分辨率的金字塔式结构, 不同尺度的小波变换刻画图像在不同分辨率条件下的特征. 这种组织结构可以有效地减少图像场景 (目标背景) 的亮度变化所产生的影响. 在近镜头球网的检测中, 我们根据二维 Mallat 算法对原始视频图像进行 2 层小波分解, 保留两次分解的低频图像用来检测球网.

5 球网检测算法

本节结合以上的叙述, 给出具体的球网检测算法, 算法流程图见下页图 3.

1) 镜头分类. 不同镜头类型的球网基元尺寸大小差异明显. 首先采用 Xu^[8] 提出的基于主颜色的镜头分类方法将镜头分为远、中、近三类.

2) 生成图像块, 估计纹理特征. 由于远镜头极少发生射门等事件, 因此只对中镜头和近镜头做球网检测. 中镜头中出现的球网通常只占图像的一部

分,我们将每帧图像分为 4×4 共 16 个等大图像块,对每个图像块估计其 5 维分形纹理特征. 对于近镜头,对整图进行 2 层小波分解. 我们知道,对大小为 $m \times n$ 的图像进行第 i 层小波分解后,图像大小将变成 $m/2^i \times n/2^i$,因此,检测时将第一层分解得到的低频图像分为 2×2 共 4 个等大图像块,第二层分解得到的低频图像不分割. 估计这 5 个图像块的 5 维纹理特征向量.

3) SVM 分类. 首先用 100 个正样本和 100 个负样本进行训练. 将每一帧产生的所有 5 维特征向量依次送入训练好的支持向量机分类,输出结果为图像块的类别,即是否是球网图像. 如果一帧中至少有一个图像块被检测为球网,那么该帧被认为是有球网出现的帧. 这种金字塔结构的检测方法对球网大小的变化具有很强的适应性,同时也在很大程度上排除了背景的干扰.

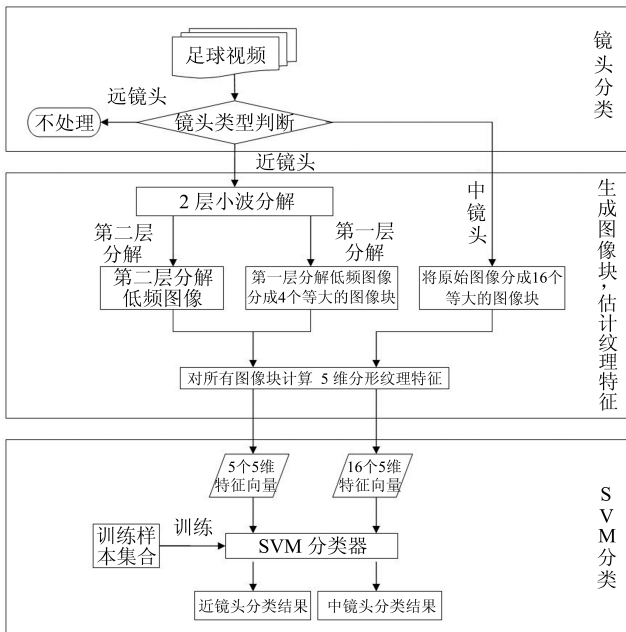


图 3 球网检测算法流程图

Fig. 3 Framework of the net detection method

6 实验结果

我们选取了有别于样本视频的四段足球视频(世界杯、德甲、英超、意甲)做测试视频,在四段不同的足球视频(世界杯、德甲、英超和意甲)上检测有球网的帧.图4为检测结果中的几个例子.由结果可以看出,在背景复杂、球网形变剧烈和基元大小不一的情况下,本文的算法仍然能够很好地检测出球网.由于其它网状物体干扰而造成错检和镜头快速移动或图像质量较差造成漏检的例子如图5所示.



图 4 检测出的有球网的帧

Fig. 4 Net frames that were detected

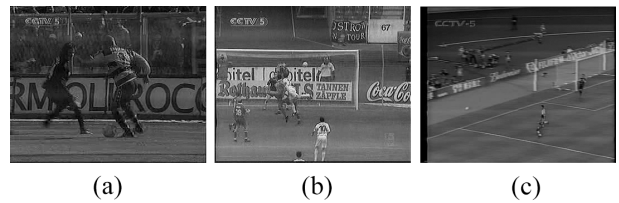


图 5 (a) 为错检的帧, (b) 和 (c) 为漏检的帧

Fig. 5 (a) An error example, (b) and (c) missing examples

因为在事件检测中,线索的抽取都是以镜头为单位^[9],因此作为上层事件检测的一个线索,需要把单帧的检测结果转换成以镜头为单位的检测结果.实验中设定了一个简单的规则,即一个镜头中,如果出现球网的帧数达到该镜头总帧数的 80% 就认为此镜头内出现球网,否则认为无球网.对四段不同足球视频的球网镜头检测结果见表 2.实验证明,球网对上层的事件检测是一个很有意义的线索.

表 2 球网镜头检测结果

Table 2 Net detection results

	世界杯	德甲	英超	意甲
镜头总数	59	47	91	55
球网镜头总数	3	4	5	4
检测结果	3	3	4	6
错检数	0	0	0	2
漏检数	0	1	1	0
检出率	100%	75%	80%	100%
准确率	100%	100%	100%	66.67%

由实验结果可以看出,世界杯比赛视频检测结果最好,达到了零错误率,这是因为世界杯比赛场地草地的颜色均一,与比赛现场其它物体的颜色区别明显,在利用主颜色进行镜头分类时能达到很高的准确率.同时,此段视频的图像质量较其它视频高很多,因此检测结果最好.漏检的情况发生于镜头的快速移动导致图像模糊,球网的网状特征被破坏,因而发生漏检.错检发生在意甲比赛视频上,原因在于该

比赛的球场和观众之间有很高的铁丝网隔离,如图5(a),有铁丝网出现的镜头被误判作球网镜头,这同时也说明此方法同样适用于一般性网状纹理的检测.

用本文提出的方法在篮球和网球比赛视频上进行球网检测,检测出的有球网的帧举例如图6.

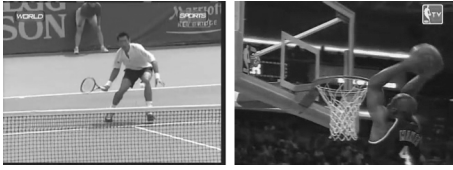


图6 在网球和篮球比赛视频中检测出的有球网的帧举例

Fig. 6 Example of detected net frames in tennis and basketball game video

7 结论

本文从纹理分析的角度出发,以足球视频中球网的检测为例,采用分形纹理特征描述网状纹理.为了达到高效快速的处理,引入信息论的观点,根据因果关系进行分形纹理特征的选择,达到了降维的目的,并减少了无关特征对分类结果的影响,很大程度上提高了运算速度.考虑到远、中、近镜头中球网的基元大小差异,采用了小波多分辨率分析方法,对近镜头图像进行不同尺度的压缩,保证球网纹理基元大小在可处理的范围内,然后对不同尺度的图像估计分形纹理特征并使用支持向量机分类.实验结果表明本文提出的网状纹理检测方法能够处理各种复杂背景的情况,并且不受基元大小和形变的影响.在未来的工作中,可以将该方法推广至其它具有网状纹理结构的物体的检测.

References

- 1 Fisher Y. *Fractal Image Compression: Theory and Application*. New York: Springer-Verlag, 1995
- 2 Mandelbrot B B. *The Fractal Geometry of Nature*. San Francisco: Freeman, 1982
- 3 Keller J M, Chen S, Crownover R M. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 1989, **45**(2): 150~166
- 4 Sarkar N, Chaudhuri B B. Efficient approach to estimate fractal dimension of textural images. *Pattern Recognition*, 1992, **25**(9): 1035~1041
- 5 Kruppa H, Schiele B. Hierarchical combination of object models using mutual information. In: Proceedings of

British Machine Vision Conference. Manchester, UK, 2001. 103~112

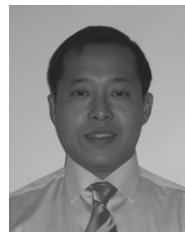
- 6 Cover T M, Thomas J A. *Elements of Information Theory*. New York: John Wiley, 1991
- 7 Mallat S G. *A Wavelet Tour of Signal Processings*. San Diego: Academic Press, 1998. 302~310
- 8 Xu P, Xie L, Chang S F, Divakaran A, Vetro A, Sun H. Algorithms and system for segmentation and structure analysis in soccer video. In: Proceedings of IEEE International Conferences on Multimedia & Expo. IEEE, 2001. 576~579
- 9 Jin G Y, Tao L M, Xu G Y. Hidden Markov model based events detection in soccer video. In: Proceedings of International Conference of Image Analysis and Recognition. Porto, Portugal, 2004. 605~612



曹媛媛 清华大学计算机系人机交互与媒体集成研究所博士研究生. 主要研究方向为计算机视觉, 多媒体信息处理. 本文通信作者. E-mail: caoyy04@mails.tsinghua.edu.cn

(CAO Yuan-Yuan Ph.D. candidate at the HCI Institute of Department of Computer Science and Technol-

ogy, Tsinghua University. Her research interest covers computer vision and multimedia information processing. Corresponding author of this paper.)



杨波 2004年获清华大学计算机科学与技术专业博士学位. 主要研究方向为计算机视觉、图像和视频的分析与检索.

(YANG Bo Received his Ph.D. degree in computer science & technology from Tsinghua University in 2004. His research interest covers computer vision, image and video analysis and re-

trieval.)



徐光祐 清华大学计算机系责任教授. 1963年清华大学自动控制系本科毕业. 目前主要研究领域为计算机视觉, 多媒体技术, 人机交互和普适计算技术.

(XU Guang-You Chair professor at Department of Computer Science & Technology, Tsinghua University. He graduated from Department of Auto-

matic Control at Tsinghua University in 1963. His research interest covers computer vision, multimedia, human computer interaction, and pervasive computing.)