

硕 士 研 究 生 读 书 报 告



题目 实例搜索中的拓扑空间验证

作者姓名 张凯

作者学号 21551016

指导教师 李启雷

学科专业 移动互联网与游戏开发

所在学院 软件学院

提交日期 二○一五年 十二月

Topological Spatial Verification for Instance Search

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: QiLei Li

By

Kai Zhang

Zhejiang University, P.R. China

2015

摘要

本文提出了一种利用图片进行实例搜索的验证方法，特别是处理一些非二维和非刚性的查询，并且这些查询往往还伴随着复杂的空间形态变化。现有实例搜索的模型是通过映射多张图片上的相似的关键点，这种模型往往是基于线性的变换。而本文将要探索的模型是在现实生活中极其常见的非线性实例变换，即通过拓扑序列关键点定位图片上非线性空间转换。文中提出了一种叫做“轮廓适配”的方案，这种方案使用一种新技术，可以用于弹性地验证拓扑空间与三角图形的一致性。简单来讲，就是通过空间拓扑的相对定位勾勒一个三角图形，让它的边缘与实例搜索中关键点的拓扑布局相匹配，接下来将估算三角图形的共同边缘的数量，以此来达到实例搜索的功能。

**关键词**：实例搜索，非二维，非线性变换，轮廓适配

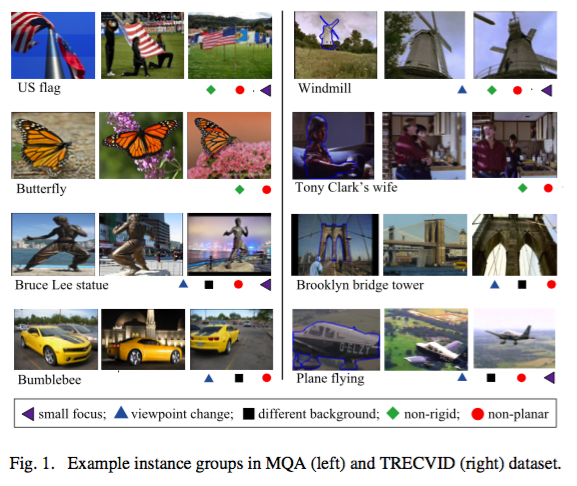
Abstract

The paper proposes an elastic spatial verification method for instance search, particularly for dealing with non-planar and non-rigid queries exhibiting complex spatial transformations. Different from existing models that map key points between images based on a linear transformation, the model this paper discussed exploits a topological arrangement of key points to address the non-linear spatial transformations that are extremely common in real life situations. In particular, the paper proposes a novel technique to elastically verify the topological spatial consistency with the triangulated graph through a “sketch-and-match” scheme. Generally specking, sketching a triangulated graph by relative positioning rather than absolute coordinating key points at first, then estimating the number of common edges between the triangulated graphs, this is what the method the paper is discussing.

**Keywords：**instance search, non-planar, non-linear transformations,

sketch-and-match

1介绍

实例搜索（INS）是一个实际的问题，这个问题首先在 TRECVID（视频检索国际权威评测）中提出，主要的是为了在一组视频集合之中检索到想要的实例。这里所说的实例是指一个可见的实体，比如说一个标志性建筑，一个地理坐标，甚至可以是一个人。概括的说，这里的搜索被定义在“实例”级别，而传统的搜索概念只是被定义在“语义”级别。对于现实中的相同实例，搜索返回的结果应该被描述为同一个实例。不同于“相似图片搜索”要求严格的相似特征，实例搜索对于相似特征的关注通常比较宽松，甚至相关的目标可以与搜索的实例一样被展示成不同的形式。特别的是，实例搜索有比较基础但是很广泛的应用，比如说档案视频的搜索，执法，个人视频的组织浏览，商标保护等等。

实例搜索的难点在于海量的搜索实例和复杂的画面条件。对于图1来说，存在以下的难点：小焦点，不同的视角，不同的背景，非二维和非刚性的目标。实例搜索中搜索目标通常是很小的查询和参考图像。因此，相同的实例往往呈现在不同的背景之中。更为困难的一点在于，SIFT（尺度不变特征变换） 通常还表现为快速的视角切换。所以，实例搜索中的“实例”可能会被呈现在非二维和非刚性的场景之中，甚至于是在不同的画面质量的3D 和非刚性的场景之中。

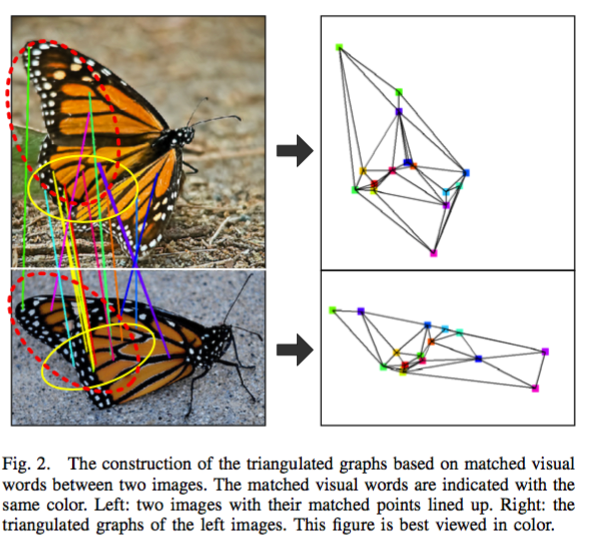
最新的实例搜索系统基于 BoW 模型，这个模型首先被应用于图像搜索。我们引入了有限的变体以修改这个模型，这样这个模型就可以用来解决前面提到关于实现实例搜索所要面对的难点。这些努力包括：探索查询和参考图片中的非对称特征，通过权衡对象和背景来构建搜索，在索引和搜索之前将多个参考图片分割成为多个小的对象目标。在实例搜索中空间验证是一个重要的组成部分，因为在要查询到目标小并且表现出复杂的空间变换的情况下，特征匹配是非常重要的一个环节。一方面来说，覆盖很小的区域的小焦点实例意味着更少的可用信息；另一方面来说，不同的相关图片上的局部特征是不稳定的，特别是通过不同的角度去看一个很小的非二维和非刚性的实例。尽管之前有人研究过图片搜索中的许多空间验证的课题，但是对于实例搜索这方面来说空间验证仍然是一个空白的领域。本文探索的就是有关实例搜索方面的空间验证技术。

2拓扑空间验证

我们首先简要地介绍了实例搜索所需要的空间配置，然后提出我们的方法来处理这些情况。最后，一些细节被附在最后，这些细节讨论了我们的基于三角图形的技术。

A.空间配置

让 R 和 Q分别表示查询图像 R 和图像 Q的齐次坐标的所有特征点。这两个平面的相关的空间坐标定位可以被一个奇次矩阵 H相关联，它们之间的关系可以表示成 Q=HR。这个公式定义了从不同视角观察图像的点对点映射。常用的仿射变换为单应，适用于近重复的搜索（缩放，旋转，裁剪）或近平面实例（地标，绘画） 。



当这个公式应用到3D对象的时候，不同视角的关系可以用一个基本的矩阵 F关联起来：QTFR=0 。然而，这个基本的矩阵只能被应用于直线上的点对点的映射，并不适用于空间验证。更加复杂的情况是那些要考虑非刚性实例验证情况，比如说图2上面的蝴蝶。之前的分析模型并不适用于这些违反了非刚性原则的情况。

为了解决这些问题，我们从另一个角度来寻找答案。尽管这些非二维和非刚性的对象没有一个统一的变换，但是这样的空间拓扑从三维物体的角度来说每个固定的视角来看都会有相对应的形状和一个刚性的结构。比如说，对于图1上面的李小龙塑像来说，如果只是偏移了一个比较小的视角，那么图像上面的相关特征点的坐标在二维平面来看可以近似地认为保持相同的位置。同样，图2 上的蝴蝶尽管表现出非刚性的特征，一些刚性结构还是会保持它们的相对空间布局结构，比如说蝴蝶的翅膀。在下一小节，我们将会提供一个弹性的空间验证的方法，这个方法能够为在空间视角变换过程中非刚性转换提供证据。

B. 轮廓适配

考虑到上一节中提到的复杂的空间配置，我们需要一个模型，这个模型在识别不一致空间布局不是很弱、在指定真实空间配置不是很严格。并且，这个模型应该能够允许小的运动和视角的变换，同时收集对于非二维和非刚性的实例的局部化匹配的证据。能够高度相似的图像，并且能够有效过滤反常的空间配置。

“拓扑”是一些关于空间的属性，这些属性描述了空间中的变形，包括拉伸、扭曲和弯曲。比如说，一个圆可以拓扑地表现成为一个椭圆。拓扑可以用于提取不同的对象之间的内在联系，当然前提是忽略它们之间的细节。本文中，我们将拓扑认为是一个很好的属性，这个属性可以用于实例验证中模型化复杂的空间配置，并且准备为空间拓扑验证准备一个叫做“轮廓适配”的策略。

我们使用Delaunay三角剖分算法（以下简称 DT 算法）来为空间拓扑简历模型。DT 算法是一种在计算机中使用的技术，这种技术用来建立点集的网络。DT 算法将点集中的点使用三角形组织起来，使用这样的方式点与点之间的边缘就会变得稳定，对于小的空间变化来说就可以近似的看成相对拓扑保持不变。对于实例搜索来说，给出要查询的图像 R 和图像 Q 之间所有匹配的轮廓，DT 算法匹配图像之间相关联的空间结构。图2展示了一个匹配图像之间的三角网状结构的例子。

我们通过将图像上面的可匹配点的集合作为输入，应用于 DT 算法来匹配三角图形。尽管这样的三角图形是一个平面的图形，但图像上面的所有边缘都被看成是非平面的图形。通过描绘图像，DT 算法能够通过限制所有三角图形中最小三角图形的总数来尽可能地避免那些过于细小的三角图形。所以，这样的拓扑布局被勾画成一个三角的图形，这样的图形可以近似的连接图形上面的点集。例如，每一条边代表空间中相近的两个点，每一个三角形代表了空间中相近的两个三角形，所有这些点集和三角形的集合组成了一个轮廓，这个轮廓组成了匹配位置的有序拓扑。通过这种方式，抛弃了绝对的位置匹配元素，取而代之的是相对的位置匹配元素用于勾画上面所说的轮廓。并且这里所说的构图方式并不依赖于图形的缩放、旋转和视角的变化。

为了使结果图形是可比较的，必须要强调一对一的结构映射。这样实现是通过 Q图上的一个点对应 R 图上面的一个点，这其中对应的关系是通过最小化的汉明码区别来。每一个点与一个32位的汉明码组成的特征码作为区分彼此的特征码。这种限制有效地避免了大量的数据造成的冗余匹配问题，这样就能够纠正图片上的重复的图案所产生的相似数据。需要注意的是如果有大量的可视化词汇，那么这些可视化的词汇的两个点之间的错误匹配被期望于最小化，这样的要求是通过强制性的一对一映射来实现的。尽管如此，如果是遇到重复的图案或者是非字符化的表面之间的比较的话，这样的方案就会失效。之前的方案就会变成武断的点与点的匹配，所以接下来简要讨论的方案将会避免这个问题，将这些重复的图案或者字符的处理减少为几个甚至是没有。

通过三角化图形之后，空间上的连贯性将会转化为图形匹配的问题。使用∆Q 表示图形 Q 的三角化图形，使用∆R 表示图形 R 的三角化图形，那么这两个图形之间的几何连贯性叫做相似因子（bonus factor），他们之间的关系可以表示成：

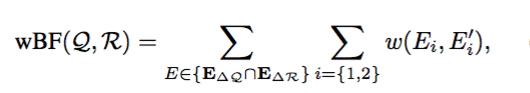
BF (Q, R) = ∥E∆Q ∩ E∆R∥

其中E∆Q表示∆Q的边集，BF 表明了图形Q 和 R 之间的共同边的数量。如果两条边的可视化元素相同的话，那么这两条边就会被认为是相同的两条边。图形R 与图形 Q的匹配度将会通过这样的相同边的条数来衡量。这样的方案在时间过程中效果相当不错，因为相互匹配的边将会被分成一组，这样的结果就是降低了错误匹配的概率。

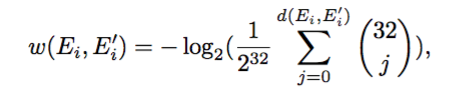
除了上面所说的图形匹配策略，我们还比较了一个可变的选项，比如说，比较两个图形之间的相同三角图形而不是相同的边。然而我们发现还是第一种方案对于我们实验的案例具有最好的效果。尽管这个方案比较简单，但是它可以高效率地给出了两个三角图形之间的强联系。值得注意得是，我们这里所说的空间验证需要在极短的时间内处理完成，而不能够执行后期处理的方式。所以我们模仿 WGC 的方式，为每一个点引入了一个局部的坐标，并将这些坐标都不放入一个局部的文件中，通过这种方式，我们可以使得搜索的时间持续更长。

C.加权函数

我们还引入了一个加权函数来更好地估计图形之间的空间联系。本质上来讲，图形之间的每一条相似的边都应该是相似因子中的一个元素。具体地讲，我们测量每个公共边缘E的强度，作为其两个端点E1和E2的匹配置信：



其中 w(Ei,Ei′) 表示的是查询图像中的点 Ei 以及这个点所对应的索引图像中的点 Ei′ ，这里的每一个点用到了我们之前提到的32位汉明码。这样，w(Ei,Ei′) 就可以表示成为基于二项式的公式：



其中的 d(Ei,Ei′) 表示这两点之间的汉明码的差值。这样的加权函数通过从0到32的汉明码之间的区别来实现。高度相似的匹配点集将会拥有更高的权值，这样这些点集的对应边集也就相应的拥有了更高的匹配度。

D.相关讨论

对于轮廓匹配的相关讨论，对于 DT 算法来说，每个图像上匹配的特征点集首先被三角化成为一个近似的空间图形。接着拓扑布局的相似性是由图形的相应的相似性决定。轮廓的确定是通过使用相对的坐标位置而不是绝对的坐标位置。图2给出了 DT 算法的一个具体实现的例子。因为蝴蝶的翅膀是非刚性的，所以说从一个匹配坐标转换到另一个坐标点的线性转换是不存在的。另外的一些技术比如说 RANSAC 只能保证匹配图像的部分碎片，图2上的任意一个蝴蝶都会被定义成刚性的。根据RANSAC技术，图2 上面黄色椭圆包括的部分里面有6个点，但是7个有效的匹配点却被排除在外。对于这个例子来说，E-WGC 算法只能定位5个真正匹配的相似点。而对于 DT 算法，能够将两个翅膀都分别识别出来，并且得出的相似程度达到了0.67，这是因为只有相对的位置坐标被使用在算法当中。除了能够定位相关的匹配之外，对象中角度偏移不大的非刚性的和非二维的部分能够被识别出来。这种实验中的假设通常会出现在实际的场景中。比如说，每个人走路的时候都只是移动每个身体部分很小的角度。在图2中0.67的相似度也包括两个蝴蝶翅膀之间的相似部分。

对于DT 算法的深入理解。DT 算法有以下的几个特点：

（1）考虑的是元素之间的相对坐标位置

（2）对于任何转换的假设都不是强制性的

（3）对于匹配的元素之间的小角度变换是可以容忍的

对于 WGC 算法来说，主要考虑的是确切的元素之间的拓扑结构，因此这种算法更多的依赖几何学上的连贯性。相比较于严格的空间验证算法，不需要事先对类型的实例和转换有所了解，因此对于几何学上的检查认证是比较宽松的。然而，因为 DT 算法允许没有转换模型的假设，所以说 DT 算法是一种弹性模型，更适用于非刚性和非二维的实例，并且这样的实例可以适应不同的摄影条件。DT 算法和其它空间验证算法的最基本的区别在于它不涉及修剪错误匹配和模型估计。相反的，DT 算法通过基于公式1和公式3的局部拓扑连续列举那些潜在的正确匹配，并且允许那些合适的匹配，这些合适的匹配不会引入任何基于标准2的预判。这样 DT 算法就致力于寻找那些正确的匹配而不去惩罚那些错误的匹配，我们把这种方式叫做红利因子。­

一些空间验证方法需要依赖局部的几何参数，比如说方向，缩放等。然而，经过我们的实验验证，这样的方法并不适合那些集合参数有可能被干扰和偏移的情况。相反的，我们的方法仅仅需要 SIFT 特征的一些局部信息，这样做可以让我们的方法更加的清晰和稳定。

估算 SIFT 特征的缩放和方向的过程中可能会遇到很多的干扰信息，即使是使用 SIFT对于平面的粗线条匹配也被认为是不容易的，对于大视角下检测这些特征就更加的困难了。基于实例搜索中的非刚性和非二维性的平面是普遍的请况，清晰的估算局部的几何参数是一个很大的挑战。所以，使用 SIFT 方法转换局部的几何参数就变得十分的不可靠。SIFT的局部几何信息也会产生严重的偏颇。

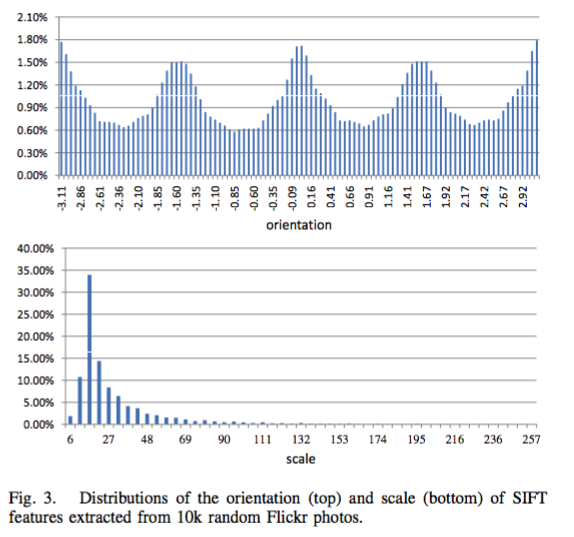


图3 绘制了从Flickr（雅虎旗下的图片分享网站）上利用Hessian矩阵仿射检测器和SIFT描述符随机抽取的1万条中得出的 SIFT特征的关于方向（上面）和缩放（下面）的统计信息。如图所示，虽然SIFT特征是由完全随机的图像取样，但是图片的局部几何形状的分布却是非常不统一的，尤其是在方向和缩放这两个参数上面存在着很大的偏差。尽管方向的偏差可以解释为人们拍照姿势的习惯，但是缩放方面严重的偏差就只能归咎于几何平面估算方面的局限了。因此，在这样大的偏差的前提下估算空间转化是存在着错误倾向的。相反的，DT 算法并不会受这些方向和缩放偏差的影响，原因是这种空间验证方法并不会去使用方向和缩放这两个因子。

E.复杂度分析

时间复杂度。DT 算法的两个主要的步骤是三角化图形和统计相同边的数量。第一步的复杂度可以使用分治法，那么它的复杂度可以表示成为 O（n \* logn），n 表示的是图形中的点的数量。第二步的复杂度度可以线性地表示成为 O（e），e 表示的是图形之间的相同的边的数量。e 同样可以使用符号 n 来表示，这样总的时间复杂度就是 O（n \* logn） 。

空间复杂度。算法对于空间的消耗主要就是存储查询图形Q 和索引图形 R 的局部匹配点集。对于一个有 N 张图形的数据集，M 是联系图形之间的最大样本匹配点集，那么至少需要4\*M\*N个整形数据。因此空间复杂度取决于 N 的大小。我们的测试用例取 M=30，N=106，最终消耗了228MB 的空间储存。对于106这样数量级别的图片所占用的空间（大约15GB） 来说，这样的空间消耗几乎可以是忽略不计的。

3总结

实例搜索中的空间验证在现今是一个非常具有挑战性的课题。对于复杂的空间验证中，太过于严格的限制或者是太过于宽松的限制都可能会导致验证的失败。本文中，我们提供了一种基于三角图形的空间验证技术方法，这种方法使用了介于严格的限制和宽松的限制之间的一种中和的方式。为了展示空间验证的配置复杂度，我们探索了空间转换过程中的拓扑透视图的变体。我们倾向于通过拓扑模型进行特征分析而不是传统的几何映射。一方面来说，DT 算法对于匹配过程中的局部变量不敏感；另一方面，这种算法对于严重的变化和出现在拓扑布局中的重复图案敏感。我们广泛的实验结果表明我们的方法对于实例搜索具有有效性和高效性。

参考文献

[1]  A. F. Smeaton, P. Over, and W. Kraaij, “Evaluation campaigns and trevid,” in ACM Multimedia Information Retrieval, 2006.

[2]  D. Lowe, “Distinctive image features from scale-invariant keypoints,” IJCV, vol. 60, no. 2, pp. 91–110, 2004.

[3]  R. Arandjelovic ́ and A. Zisserman, “Three things everyone should know to improve object retrieval,” in IEEE Conference on Computer Vision and Pattern Recognition, 2012.

[4]  D. Le, C. Zhu, S. Poullot, and S. Satoh, “National institute of informat- ics, Japan at TRECVID 2011,” in TRECVID, 2011.

[5]  W. Zhang, C.-C. Tan, S.-A. Zhu, T. Yao, L. Pang, and C.-W. Ngo, “Vireo @ trecvid 2012: Searching with topology, recounting will small concepts,learning with free examples,” in TRECVID, 2012.

[6]  Z. Zhang, R. Albatal, C. Gurrin, and A. F. Smeaton, “Trecvid 2013 experiments at Dublin City University,” in TRECVID, 2013.

[7]  J. Sivic and A. Zisserman, “Video Google: A text retrieval approach to object matching in videos,” in International Conference on Computer Vision, vol. 2, Oct. 2003, pp. 1470–1477.

[8] C.-Z. Zhu, H. Je ́gou, and S. Satoh, “Query-adaptive asymmetrical dissimilarities for visual object retrieval,” in International Conference on Computer Vision, Oct. 2013.

[9] W. Zhang and C.-W. Ngo, “Searching visual instances with topology checking and context modeling,” in ACM International Conference on Multimedia Retrieval, 2013.

[10] R. Tao, E. Gavves, C. G. M. Snoek, and A. W. M. Smeulders, “Locality in generic instance search from one example,” in IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[11] C. G. M. Snoek, K. van de Sande, A. Habibian, S. Kordumova, Z. Li, M. Mazloom, S. Pintea, R. Tao, D. Koelma, and A. W. M. Smeulders, “The MediaMill at trecvid 2013: : Searching concepts, objects, instances and events in video,” in TRECVID, 2013.

[12] Z. Li, E. Gavves, K. E. A. van de Sande, C. G. M. Snoek, and A. W. M. Smeulders, “Codemaps segment, classify and search objects locally,” in IEEE International Conference on Computer Vision, 2013.

[13] H. Je ́gou, M. Douze, and C. Schmid, “Hamming embedding and weak geometric consistency for large scale image search,” in European Conference on Computer Vision, 2008.

[14] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, “Object retrieval with large vocabularies and fast spatial matching,” in IEEE Conference on Computer Vision and Pattern Recognition, 2007.

[15] W. Zhao, X. Wu, and C. W. Ngo, “On the annotation of web videos by efficient near-duplicate search,” IEEE Transactions. on Multimedia, vol. 12, no. 5, pp. 448–461, 2010.

[16] Y.Zhang,Z.Jia,andT.Chen,“Imageretrievalwithgeometrypreserving visual phrases,” in IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[17] B. Fernando, E. Fromont, and T. Tuytelaars, “Mining mid-level features for image classification,” International Journal of Computer Vision, vol. 108, no. 3, pp. 186–203, 2014.

[18] W. Zhang, L. Pang, and C.-W. Ngo, “Snap-and-ask: Answering multi- modal question by naming visual instance,” in ACM Multimedia, 2012.

[19] R. I. Hartley and A. Zisserman, Multiple View Geometry in Computer

Vision. Cambridge University Press, ISBN: 0521540518, 2004.

[20] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” Communications of the ACM, vol. 24, no. 6, pp. 381–395,

1981.

[21] Y. Avrithis and G. Tolias, “Hough pyramid matching: Speeded-up ge-

ometry re-ranking for large scale image retrieval,” International Journal

of Computer Vision, vol. 107, no. 1, pp. 1–19, 2014.

[22] X. Wu, A. G. Hauptmann, and C.-W. Ngo, “Practical elimination of near-duplicates from web video search,” in ACM Multimedia, 2007, pp.

218–227.

[23] O.Chum,J.Philbin,M.Isard,andA.Zisserman,“Scalablenearidentical

image and shot detection,” in International Conference on Image and

Video Retrieval, 2007.

[24] R. Arandjelovic and A. Zisserman, “Efficient image retrieval for 3d

structures,” in The British Machine Vision Conference, 2010.

[25] O. Chum and J. Matas, “Large-scale discovery of spatially related im- ages,” IEEE Transactions on Pattern Analysis and Machine Intelligence,

vol. 32, pp. 371–377, 2010.

[26] Z. Wu, Q. Ke, M. Isard, and J. Sun, “Bundling features for large scale

partial-duplicate web image search,” in IEEE Conference on Computer

Vision and Pattern Recognition, Aug. 2009, pp. 25–32.

[27] http://mathworld.wolfram.com/Topology.html.

[28]  B. N. Delaunay, “Sur la sphe`re vide,” Bulletin of Academy of Sciences of the USSR, no. 6, pp. 793–800, 1934.

[29]  Y. Kalantidis, L. Pueyo, M. Trevisiol, R. van Zwol, and Y. Avrithis, “Scalable triangulation-based logo recognition,” in ACM International Conference on Multimedia Retrieval, April 2011.

[30]  H.Je ́gou,M.Douze,andC.Schmid,“Improvingbag-of-featuresfor large scale image search,” International Journal of Computer Vision, vol. 87, no. 3, pp. 192–212, May 2010.

[31]  H. Je ́gou, M. Douze, and C. Schmid, “On the burstiness of visual elements,” in IEEE Conference on Computer Vision and Pattern Recog- nition, 2009.

[32]  K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, “A comparison of affine region detectors,” Int. J. Comput. Vision, vol. 65, no. 1-2, pp. 43–72, Nov. 2005.

[33]  K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” IEEE Transactions on Pattern Analysis and Machine Intel- ligence, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[34]  G. Leach, “Improving worst-case optimal delaunay triangulation algo- rithms,” in In 4th Canadian Conference on Computational Geometry, 1992.

[35]  J. Philbin, M. Isard, J. Sivic, and A. Zisserman, “Lost in quantization: Improving particular object retrieval in large scale image databases,” in IEEE Conference on Computer Vision and Pattern Recognition, 2008.

[36]  D. Nister and H. Stewenius, “Scalable recognition with a vocabulary tree,” in IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 2161–2168.

[37]  Y.-G. Jiang, J. Yang, C.-W. Ngo, and A. G. Hauptmann, “Representa- tions of keypoint-based semantic concept detection: A comprehensive study,” IEEE Transactions on Multimedia, vol. 12, pp. 42–53, 2010.

[38]  C. G. M. Snoek, K. van de Sande, A. Habibian, S. Kordumova, Z. Li, M. Mazloom, S. Pintea, R. Tao, D. Koelma, and A. W. M. Smeulders, “The MediaMill trecvid 2012 semantic video search engine,” in TRECVID, 2012.

[39]  Y. Peng, X. Zhai, J. Zhang, L. Huang, N. Li, P. Tang, X. Huang, and Y. Zhao, “PKU ICST at trecvid2013 : Instance search task,” in TRECVID, 2013.

[40]  D.-D. Le, C.-Z. Zhu, and S. Satoh, “National institute of informatics, japan at trecvid 2013,” in TRECVID Workshop, 2013.