
图像处理与计算机视觉

基础，经典以及最近发展

By *xdyang* (杨晓冬 xdyang.ustc@gmail.com)

一、绪论

1. 为什么要写这篇文章

从 2002 年到现在，接触图像快十年了。虽然没有做出什么很出色的工作，不过在这个领域摸爬滚打了十年之后，发现自己对图像处理和计算机视觉的感情越来越深厚。下班之后看看相关的书籍和文献是一件很惬意的事情。平常的一大业余爱好就是收集一些相关的文章，尤其是经典的文章，到现在我的电脑里面已经有了几十 G 的文章。写这个文档的想法源于我前一段时间整理文献时的一个突发奇想，既然有这个多文献，何不整理出其中的经典，抓住重点来阅读，同时也可以共享给大家。于是当时即兴写了一个《图像处理与计算机视觉中的经典论文》。现在来看，那个文档写得很一般，所共享的论文也非常之有限。就算如此，还是得到了一些网友的夸奖，心里感激不尽。因此，一直想下定决心把这个工作给完善，力求做到尽量全面。

本文是对现有的图像处理和计算机视觉的经典书籍（后面会有推荐）的一个补充。一般的图像处理书籍都是介绍性的介绍某个方法，在每个领域内都会引用几十上百篇参考文献。有时候想深入研究这个领域的时候却发现文献太多，不知如何选择。但实际上在每个领域都有那么三五篇抑或更多是非读不可的经典文献。这些文献除了提出了很经典的算法，同时他们的 Introduction 和 Related work 也是对所在的领域很好的总结。读通了这几篇文献也就等于深入了解了这个领域，比单纯的看书收获要多很多。写本文的目的就是想把自己所了解到的各个领域的经典文章整理出来，不用迷失在参考文献的汪洋大海里。

2. 图像处理和计算机视觉的分类

按照当前流行的分类方法，可以分为以下三部分：

图像处理：对输入的图像做某种变换，输出仍然是图像，基本不涉及或者很少涉及图像内容的分析。比较典型的有图像变换，图像增强，图像去噪，图像压缩，图像恢复，二值图像处理等等。基于阈值的图像分割也属于图像处理的范畴。一般处理的是单幅图像。

图像分析：对图像的内容进行分析，提取有意义的特征，以便于后续的处理。处理的仍然是单幅图像。

计算机视觉: 对图像分析得到的特征进行分析, 提取场景的语义表示, 让计算机具有人眼和人脑的能力。这时处理的是多幅图像或者序列图像, 当然也包括部分单幅图像。

关于图像处理, 图像分析和计算机视觉的划分并没有一个很统一的标准。一般的来说, 图像处理的书籍总会或多或少的介绍一些图像分析和计算机视觉的知识, 比如冈萨雷斯的数字图像处理。而计算机视觉的书籍基本上都会包括图像处理和图像分析, 只是不会介绍的太详细。其实图像处理, 图像分析和计算机视觉都可以纳入到计算机视觉的范畴: 图像处理->低层视觉 (low level vision), 图像分析->中间层视觉 (middle level vision), 计算机视觉->高层视觉 (high level vision)。这是一般的计算机视觉或者机器视觉的划分方法。在本文中, 仍然按照传统的方法把这个领域划分为图像处理, 图像分析和计算机视觉。

3. 图像处理和计算机视觉开源库以及编程语言选择

目前在图像处理中有两种最重要的语言: c/c++和 matlab。它们各有优点: c/c++比较适合大型的工程, 效率较高, 而且容易转成硬件语言, 是工业界的默认语言之一。而 matlab 实现起来比较方便, 适用于算法的快速验证, 而且 matlab 有成熟的工具箱可以使用, 比如图像处理工具箱, 信号处理工具箱。它们有一个共同的特点: 开源的资源非常多。在学术界 matlab 使用的非常多, 很多作者给出的源代码都是 matlab 版本。最近由于 OpenCV 的兴起和不断完善, c/c++在图像处理中的作用越来越大。总的来说, c/c++和 matlab 都必须掌握, 最好是精通, 当然侧重在 c/c++上对找工作会有很大帮助。

至于开源库, 个人非常推荐 OpenCV, 主要有以下原因:

(1) 简单易入手。OpenCV 进入 OpenCV2.x 的时代后, 使用起来越来越简单, 接口越来越傻瓜化, 越来越 matlab 化。只要会 imread, imwrite, imshow 和了解 Mat 的基本操作就可以开始入手了。

(2) OpenCV 有一堆图像处理和计算机视觉的大牛在维护, bug 在逐步减少, 每个新的版本都会带来不同的惊喜。而且它已经或者逐步在移植到不懂的平台, 并提供了对 Python 的很好的支持。

(3) 在 OpenCV 上可以尝试各种最新以及成熟的技术, 而不需要自己从头去写, 比如人脸检测 (Harr, LBP), DPM (Latent SVM), 高斯背景模型, 特征检测, 聚类, hough 变换等等。而且它还支持各种机器学习方法 (SVM, NN, KNN, 决策树, Boosting 等), 使用起来很简单。

(4) 文档内容丰富, 并且给出了很多示例程序。当然也有一些地方文档描述不清楚, 不过看看代码就很清楚了。

(5) 完全开源。可以从中间提取出任何需要的算法。

(6) 从学校出来后, 除极少数会继续在学术圈里, 大部分还是要进入工业界。现在在工业界, c/c++仍是主流, 很多公司都会优先考虑熟悉或者精通 OpenCV 的。事实上, 在学术界, 现在 OpenCV 也大有取代 matlab 之势。以前的

demo 或者 source code, 很多作者都愿意给出 matlab 版本的, 然后别人再呼哧呼哧改成 c 版本的。现在作者干脆给出 c/c++ 版本, 或者自己集成到 OpenCV 中去, 这样能快速提升自己的影响力。

如果想在图像处理和计算机视觉界有比较深入的研究, 并且以后打算进入这个领域工作的话, 建议把 OpenCV 作为自己的主攻方向。如果找工作的时候敢号称自己精通 OpenCV 的话, 肯定可以找到一份满意的工作。

4. 本文的特点和结构, 以及适合的对象

本文面向的对象是即将进入或者刚刚进入图像处理和计算机视觉领域的童鞋, 可以在阅读书籍的同时参阅这些文献, 能对书中提到的算法有比较深刻的理解。由于本文涉及到的范围比较广, 如果能对计算机视觉的资深从业者也有一定的帮助, 我将倍感欣慰。为了不至太误人子弟, 每一篇文章都或多或少的看了一下, 最不济也看了摘要(这句话实在整理之前写的, 实际上由于精力有限, 好多文献都只是大概扫了一眼, 然后看了看 google 的引用数, 一般在 1000 以上就放上来了, 把这些文章细细品味一遍也是我近一两年之内的目标)。在成文的过程中, 我本人也受益匪浅, 希望能对大家也有所帮助。

由于个人精力和视野的关系, 有一些我未涉足过的领域不敢斗胆推荐, 只是列出了一些引用率比较高的文章, 比如摄像机标定和立体视觉。不过将来, 由于工作或者其他原因, 这些领域也会接触到, 我会逐步增减这些领域的文章。尽管如此, 仍然会有疏漏, 忘见谅。同时文章的挑选也夹带了一些个人的喜好, 比如我个人比较喜欢 low level 方向的, 尤其是 IJCV 和 PAMI 上面的文章, 因此这方面也稍微多点, 希望不要引起您的反感。如果有什么意见或者建议, 欢迎 mail 我。文章和资源我都会在我的 csdn blog 和 sina ishare 同步更新。在此申明: 这些论文的版权归作者及其出版商所有, 请勿用于商业目的。

个人 blog:

<http://blog.csdn.net/dcraw>

新浪 iask 地址:

<http://iask.sina.com.cn/u/2252291285/ish?folderid=868438>

本文的安排如下。第一部分是绪论。第二部分是图像处理中所需要用到的理论基础, 主要是这个领域所涉及到的一些比较好的参考书籍。第三部分是计算机视觉中所涉及到的信号处理和模式识别文章。由于图像处理与图像分析太难区分了, 第四部分集中讨论了它们。第五部分是计算机视觉部分。最后是小结。

二、 图像处理与计算机视觉相关的书籍

1. 数学

我们所说的图像处理实际上就是数字图像处理,是把真实世界中的连续三维随机信号投影到传感器的二维平面上,采样并量化后得到二维矩阵。数字图像处理就是二维矩阵的处理,而从二维图像中恢复出三维场景就是计算机视觉的主要任务之一。这里面就涉及到了图像处理所涉及到的三个重要属性:连续性,二维矩阵,随机性。所对应的数学知识是高等数学(微积分),线性代数(矩阵论),概率论和随机过程。这三门课也是考研数学的三个组成部分,构成了图像处理和计算机视觉最基础的数学基础。如果想要更进一步,就要到网上搜搜林达华推荐的数学书目了。

2. 信号处理

图像处理其实就是二维和三维信号处理,而处理的信号又有一定的随机性,因此经典信号处理和随机信号处理都是图像处理和计算机视觉中必备的理论基础。

2.1 经典信号处理

信号与系统(第2版) Alan V. Oppenheim 等著 刘树棠译

离散时间信号处理(第2版) A. V. 奥本海姆等著 刘树棠译

数字信号处理:理论算法与实现 胡广书(编者)

2.2 随机信号处理

现代信号处理 张贤达著

统计信号处理基础:估计与检测理论 Steven M. Kay 等著 罗鹏飞等译

自适应滤波器原理(第4版) Simon Haykin 著 郑宝玉等译

2.3 小波变换

信号处理的小波导引:稀疏方法(原书第3版) tephane Malla 著,戴道清等译

2.4 信息论

信息论基础(原书第2版) Thomas M. Cover 等著 阮吉寿等译

3. 模式识别

Pattern Recognition and Machine Learning Bishop, Christopher M. Springer

模式识别(英文版)(第4版) 西奥多里德斯著

Pattern Classification (2nd Edition) Richard O. Duda 等著

Statistical Pattern Recognition, 3rd Edition Andrew R. Webb 等著

模式识别(第3版) 张学工著

4. 图像处理与计算机视觉的书籍推荐

图像处理, 分析与机器视觉 第三版 Sonka 等著 艾海舟等译

Image Processing, Analysis and Machine Vision

这本书是图像处理与计算机视觉里面比较全的一本书了,几乎涵盖了图像视觉领域的各个方面。中文版的个人感觉也还可以,值得一看。

数字图像处理 第三版 冈萨雷斯等著

Digital Image Processing

数字图像处理永远的经典,现在已经出到了第三版,相当给力。我的导师曾经说过,这本书写的很优美,对写英文论文也很有帮助,建议购买英文版的。

计算机视觉: 理论与算法 Richard Szeliski 著

Computer Vision: Theory and Algorithm

微软的 Szeliski 写的一本最新的计算机视觉著作。内容非常丰富,尤其包括了作者的研究兴趣,比如一般的书里面都没有的 Image Stitching 和 Image Matting 等。这也从另一个侧面说明这本书的通用性不如 Sonka 的那本。不过作者开放了这本书的电子版,可以有选择性的阅读。

<http://szeliski.org/Book/>

Multiple View Geometry in Computer Vision 第二版 Harley 等著

引用达一万多次的经典书籍了。第二版到处都有电子版的。第一版曾出过中文版的,后来绝版了。网上也可以找到中英文版的电子版。

计算机视觉: 一种现代方法 DA Forsyth 等著

Computer Vision: A Modern Approach

MIT 的经典教材。虽然已经过去十年了，还是值得一读。期待第二版

Machine vision: theory, algorithms, practicalities 第三版 Davies
著

为数不多的英国人写的书，偏向于工业应用。

数字图像处理 第四版 Pratt 著

Digital Image Processing

写作风格独树一帜，也是图像处理领域很不错的一本书。网上也可以找到非常清晰的电子版。

5. 小结

罗嗦了这么多，实际上就是几个建议：

(1) 基础书千万不可以扔，也不能低价处理给同学或者师弟师妹。不然到时候还得一本本从书店再买回来的。钱是一方面的问题，对着全新的书看完全没有看自己当年上过的课本有感觉。

(2) 遇到有相关的课，果断选修或者蹭之，比如随机过程，小波分析，模式识别，机器学习，数据挖掘，现代信号处理甚至泛函。多一些理论积累对将来科研和工作都有好处。

(3) 资金允许的话可以多囤一些经典的书，有的时候从牙缝里面省一点都可以买一本好书。不过千万不要像我一样只囤不看。

三、 计算机视觉中的信号处理与模式识别

从本章开始，进入本文的核心章节。一共分三章，分别讲述信号处理与模式识别，图像处理与分析以及计算机视觉。与其说是讲述，不如说是一些经典文章的罗列以及自己的简单点评。与前一个版本不同的是，这次把所有的文章按类别归了类，并且增加了很多文献。分类的时候并没有按照传统的分类方法，而是划分成了一个小的门类，比如 SIFT, Harris 都作为了一类，虽然它们都可以划分到特征提取里面去。这样做的目的是希望能突出这些比较实用且比较流行的方法。为了以后维护的方便，按照字母顺序排的序。

1. [Boosting](#)

Boosting 是最近十来年来最成功的一种模式识别方法之一，个人认为可以和 SVM 并称为模式识别双子星。它真正实现了“三个臭皮匠，赛过诸葛亮”。只要保证每个基本分类器的正确率超过 50%，就可以实现组合成任意精度的分类器。这样就可以使用最简单的线性分类器。Boosting 在计算机视觉中的最成功的应用无疑就是 Viola-Jones 提出的基于 Haar 特征的人脸检测方案。听起来似乎不可思议，但 Haar+Adaboost 确实在人脸检测上取得了巨大的成功，已经成了工业界的事实标准，并且逐步推广到其他物体的检测。

Rainer Lienhart 在 2002 ICIP 发表的这篇文章是 Haar+Adaboost 的最好的扩展，他把原始的两个方向的 Haar 特征扩展到了四个方向，他本人是 OpenCV 积极的参与者。现在 OpenCV 的库里面实现的 Cascade Classification 就包含了他的方法。这也说明了盛会（如 ICIP, ICPR, ICASSP）也有好文章啊，只要用心去发掘。

[1997] A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting

[1998] Boosting the margin A new explanation for the effectiveness of voting methods

[2002 ICIP TR] Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection

[2003] The Boosting Approach to Machine Learning An Overview

[2004 IJCV] Robust Real-time Face Detection

2. [Clustering](#)

聚类主要有 K 均值聚类，谱聚类和模糊聚类。在聚类的时候如果自动确定聚类中心的数目是一个一直没有解决的问题。不过这也很正常，评价标准不同，得到的聚类中心数目也不一样。不过这方面还是有一些可以参考的文献，在使用的時候可以基于这些方法设计自己的准则。关于聚类，一般的模式识别书籍都介绍的比较详细，不过关于 cluster validity 讲的比较少，可以参考下面的文章看看。

[1989 PAMI] Unsupervised Optimal Fuzzy Clustering

[1991 PAMI] A validity measure for fuzzy clustering

[1995 PAMI] On cluster validity for the fuzzy c-means model

[1998] Some New Indexes of Cluster Validity

[1999 ACM] Data Clustering A Review

[1999 JIIS] On Clustering Validation Techniques

[2001] Estimating the number of clusters in a dataset via the Gap statistic

[2001 NIPS] On Spectral Clustering

[2002] A stability based method for discovering structure in clustered data

[2007] A tutorial on spectral clustering

3. [Compressive Sensing](#)

最近大红大紫的压缩感知理论。

[2006 TIT] Compressed Sensing

[2008 SPM] An Introduction to Compressive Sampling

[2011 TSP] Structured Compressed Sensing From Theory to Applications

4. [Decision Trees](#)

对决策树感兴趣的同学这篇文章是非看不可的了。

[1986] Introduction to Decision Trees

5. [Dynamical Programming](#)

动态规划也是一个比较使用的方法，这里挑选了一篇 PAMI 的文章以及一篇 Book Chapter

[1990 PAMI] using dynamic programming for solving variational problems in vision

[Book Chapter] Dynamic Programming

6. [Expectation Maximization](#)

EM 是计算机视觉中非常常见的一种方法，尤其是对参数的估计和拟合，比如高斯混合模型。EM 和 GMM 在 Bishop 的 PRML 里单独的作为一章，讲的很不错。关于 EM 的 tutorial，网上也可以搜到很多。

[1977] Maximum likelihood from incomplete data via the EM algorithm

[1996 SPM] The Expectation-Maximization Algorithm

7. [Graphical Models](#)

伯克利的乔丹大师的 Graphical Model，可以配合这 Bishop 的 PRML 一起看。

[1999 ML] An Introduction to Variational Methods for Graphical Models

8. [Hidden Markov Model](#)

HMM 在语音识别中发挥着巨大的作用。在信号处理和图像处理中也有一定的应用。最早接触它是跟小波和检索相关的，用 HMM 来描述小波系数之间的相互关系，并用来做检索。这里提供一篇 1989 年的经典综述，几篇 HMM 在小波，分割，检索和纹理上的应用以及一本比较早的中文电子书，现在也不知道作者是谁，在这里对作者表示感谢。

[1989] A tutorial on hidden markov models and selected applications in speech recognition

[1998 TSP] Wavelet-based statistical signal processing using hidden Markov models

[2001 TIP] Multiscale image segmentation using wavelet-domain hidden Markov models

[2002 TMM] Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models

[2003 TIP] Wavelet-based texture analysis and synthesis using hidden Markov models

Hmm Chinese book.pdf

9. [Independent Component Analysis](#)

同 PCA 一样，独立成分分析在计算机视觉中也发挥着重要的作用。这里介绍两篇综述性的文章，最后一篇是第二篇的 TR 版本，内容差不多，但比较清楚一些。

[1999] Independent Component Analysis A Tutorial

[2000 NN] Independent component analysis algorithms and applications

[2000] Independent Component Analysis Algorithms and Applications

10. [Information Theory](#)

计算机视觉中的信息论。这方面有一本很不错的书 Information Theory in Computer Vision and Pattern Recognition。这本书有电子版，如果需要用到的话，也可以参考这本书。

[1995 NC] An Information-Maximization Approach to Blind Separation and Blind Deconvolution

[2010] An information theory perspective on computational vision

11. [Kalman Filter](#)

这个话题在张贤达老师的现代信号处理里面讲的比较深入,还给出了一个有趣的例子。这里列出了 Kalman 的最早的论文以及几篇综述,还有 Unscented Kalman Filter。同时也有了一篇 Kalman Filter 在跟踪中的应用以及两本电子书。

[1960 Kalman] A New Approach to Linear Filtering and Prediction Problems Kalman

[1970] Least-squares estimation_from Gauss to Kalman

[1997 SPIE] A New Extension of the Kalman Filter to Nonlinear System

[2000] The Unscented Kalman Filter for Nonlinear Estimation

[2001 Siggraph] An Introduction to the Kalman Filter_full

[2003] A Study of the Kalman Filter applied to Visual Tracking

12. [Pattern Recognition and Machine Learning](#)

模式识别名气比较大的几篇综述

[2000 PAMI] Statistical pattern recognition a review

[2004 CSVT] An Introduction to Biometric Recognition

[2010 SPM] Machine Learning in Medical Imaging

13. [Principal Component Analysis](#)

著名的 PCA, 在特征的和特征降维上非常有用。

[2001 PAMI] PCA versus LDA

[2001] Nonlinear component analysis as a kernel eigenvalue problem

[2002] A Tutorial on Principal Component Analysis

[2009] A Tutorial on Principal Component Analysis

[2011] Robust Principal Component Analysis

[Book Chapter] Singular Value Decomposition and Principal Component Analysis

14. [Random Forest](#)

随机森林

[2001 ML] Random Forests

15. [RANSAC](#)

随机抽样一致性方法，与传统的最小均方误差等完全是两个路子。在 Sonka 的书里面也有提到。

[2009 BMVC] Performance Evaluation of RANSAC Family

16. [Singular Value Decomposition](#)

对于非方阵来说，就是 SVD 发挥作用的时刻了。一般的模式识别书都会介绍到 SVD。这里列出了 K-SVD 以及一篇 Book Chapter

[2006 TSP] K-SVD An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation

[Book Chapter] Singular Value Decomposition and Principal Component Analysis

17. [Sparse Representation](#)

这里主要是 Proceeding of IEEE 上的几篇文章

[2009 PAMI] Robust Face Recognition via Sparse Representation

[2009 PIIEEE] Image Decomposition and Separation Using Sparse Representations An Overview

[2010 PIIEEE] Dictionaries for Sparse Representation Modeling

[2010 PIIEEE] It's All About the Data

[2010 PIIEEE] Matrix Completion With Noise

[2010 PIEEE] On the Role of Sparse and Redundant Representations in Image Processing

[2010 PIEEE] Sparse Representation for Computer Vision and Pattern Recognition

[2011 SPM] Directionary Learning

18. [Support Vector Machines](#)

[1998] A Tutorial on Support Vector Machines for Pattern Recognition

[2004] LIBSVM A Library for Support Vector Machines

19. [Wavelet](#)

在小波变换之前，时频分析的工具只有傅立叶变换。众所周知，傅立叶变换在时域没有分辨率，不能捕捉局部频域信息。虽然短时傅立叶变换克服了这个缺点，但只能刻画恒定窗口的频率特性，并且不能很好的扩展到二维。小波变换的出现很好的解决了时频分析的问题，作为一种多分辨率分析工具，在图像处理中得到了极大的发展和应用。在小波变换的发展过程中，有几个人是不得不提的，Mallat, Daubechies, Vetteri, M.N.Do, Swelden, Donoho。Mallat 和 Daubechies 奠定了第一代小波的框架，他们的著作更是小波变换的必读之作，相对来说，小波十讲太偏数学了，比较难懂。而 Mallat 的信号处理的小波导引更偏应用一点。Swelden 提出了第二代小波，使小波变换能够快速方便的实现，他的功劳有点类似于 FFT。而 Donoho, Vetteri, Mallat 及其学生们提出了 Ridgelet, Curvelet, Bandelet, Contourlet 等几何小波变换，让小波变换有了方向性，更便于压缩，去噪等任务。尤其要提的是 M.N.Do，他是一个越南人，得过 IMO 的银牌，在这个领域著作颇丰。我们国家每年都有 5 个左右的 IMO 金牌，希望也有一两个进入这个领域，能够也让我们也敬仰一下。而不是一股脑的都进入金融，管理这种跟数学没有多大关系的行业，呵呵。很希望能看到中国的陶哲轩，中国的 M.N.Do。

说到小波，就不得不提 JPEG2000。在 JPEG2000 中使用了 Swelden 和 Daubechies 提出的用提升算法实现的 9/7 小波和 5/3 小波。如果对比 JPEG 和 JPEG2000，就会发现 JPEG2000 比 JPEG 在性能方面有太多的提升。本来我以为 JPEG2000 的普及只是时间的问题。但现在看来，这个想法太 Naive 了。现在已经过去十几年了，JPEG2000 依然没有任何出头的迹象。不得不说，工业界的惯性力量太强大了。如果以前的东西没有什么硬伤的话，想改变太难了。不巧的是，JPEG2000 的种种优点在最近的硬件上已经有了很大的提升。压缩率？现在动辄 1T, 2T 的硬盘，没人太在意压缩率。渐进传输？现在的网速包括无线传输的速度已经相当快了，渐进传输也不是什么优势。感觉现在做图像压缩越来越没有前

途了，从最近的会议和期刊文档也可以看出这个趋势。不管怎么说，JPEG2000的 Overview 还是可以看看的。

[1989 PAMI] A theory for multiresolution signal decomposition__the wavelet representation

[1996 PAMI] Image Representation using 2D Gabor Wavelet

[1998] FACTORING WAVELET TRANSFORMS INTO LIFTING STEPS

[1998] The Lifting Scheme_ A Construction Of Second Generation Wavelets

[2000 TCE] The JPEG2000 still image coding system_ an overview

[2002 TIP] The curvelet transform for image denoising

[2003 TIP] Gray and color image contrast enhancement by the curvelet transform

[2003 TIP] Mathematical Properties of the jpeg2000 wavelet filters

[2003 TIP] The finite ridgelet transform for image representation

[2005 TIP] Sparse Geometric Image Representations With Bandelets

[2005 TIP] The Contourlet Transform_ An Efficient Directional Multiresolution Image Representation

[2010 SPM] The Curvelet Transform

四、 图像处理与分析

本章主要讨论图像处理与分析。虽然后面计算机视觉部分的有些内容比如特征提取等也可以归结到图像分析中来，但鉴于它们与计算机视觉的紧密联系，以及它们的出处，没有把它们纳入到图像处理与分析中来。同样，这里面也有一些也可以划归到计算机视觉中去。这都不重要，只要知道有这么个方法，能为自己所用，或者从中得到灵感，这就够了。

1. [Bilateral Filter](#)

Bilateral Filter 俗称双边滤波器是一种简单实用的具有保持边缘作用的平缓滤波器，由 Tomasi 等在 1998 年提出。它现在已经发挥着重大作用，尤其是在 HDR 领域。

[1998 ICCV] Bilateral Filtering for Gray and Color Images

[2008 TIP] Adaptive Bilateral Filter for Sharpness Enhancement and Noise Removal

2. [Color](#)

如果对颜色的形成有一定的了解，能比较深刻的理解一些算法。这方面推荐冈萨雷斯的数字图像处理中的相关章节以及 Sharma 在 Digital Color Imaging Handbook 中的第一章“Color fundamentals for digital imaging”。跟颜色相关的知识包括 Gamma，颜色空间转换，颜色索引以及肤色模型等，这其中也包括著名的 EMD。

[1991 IJCV] Color Indexing

[2000 IJCV] The Earth Mover's Distance as a Metric for Image Retrieval

[2001 PAMI] Color invariance

[2002 IJCV] Statistical Color Models with Application to Skin Detection

[2003] A review of RGB color spaces

[2007 PR] A survey of skin-color modeling and detection methods

Gamma. pdf

GammaFAQ. pdf

3. [Compression and Encoding](#)

个人以为图像压缩编码并不是当前很热的一个话题，原因前面已经提到过。这里可以看看一篇对编码方面的展望文章

[2005 IEEE] Trends and perspectives in image and video coding

4. Contrast Enhancement

对比度增强一直是图像处理中的一个恒久话题,一般来说都是基于直方图的,比如直方图均衡化。冈萨雷斯的书里面对这个话题讲的比较透彻。这里推荐几篇个人认为不错的文章。

[2002 IJCV] Vision and the Atmosphere

[2003 TIP] Gray and color image contrast enhancement by the curvelet transform

[2006 TIP] Gray-level grouping (GLG) an automatic method for optimized image contrast enhancement-part II

[2006 TIP] Gray-level grouping (GLG) an automatic method for optimized image contrast Enhancement-part I

[2007 TIP] Transform Coefficient Histogram-Based Image Enhancement Algorithms Using Contrast Entropy

[2009 TIP] A Histogram Modification Framework and Its Application for Image Contrast Enhancement

5. Deblur (Restoration)

图像恢复或者图像去模糊一直是一个非常难的问题,尤其是盲图像恢复。港中文的 jiaya jia 老师在这方面做的不错,他在主页也给出了可执行文件。这方面的内容也建议看冈萨雷斯的书。这里列出了几篇口碑比较好的文献,包括古老的 Richardson-Lucy 方法,几篇盲图像恢复的综述以及最近的几篇文章,尤以 Fergus 和 Jiaya Jia 的为经典。

[1972] Bayesian-Based Iterative Method of Image Restoration

[1974] an iterative technique for the rectification of observed distributions

[1990 IEEE] Iterative methods for image deblurring

[1996 SPM] Blind Image Deconvolution

[1997 SPM] Digital image restoration

[2005] Digital Image Reconstruction - Deblurring and Denoising

[2006 Siggraph] Removing Camera Shake from a Single Photograph

[2008 Siggraph] High-quality Motion Deblurring from a Single Image

[2011 PAMI] Richardson-Lucy Deblurring for Scenes under a Projective Motion Path

6. [Dehazing and Defog](#)

严格来说去雾化也算是图像对比度增强的一种。这方面最近比较好的工作就是 He kaiming 等提出的 Dark Channel 方法。这篇论文也获得了 2009 的 CVPR 最佳论文奖。2 这位 003 年的广东高考状元已经于 2011 年从港中文博士毕业加入 MSRA (估计当时也就二十五六岁吧), 相当了不起。

[2008 Siggraph] Single Image Dehazing

[2009 CVPR] Single Image Haze Removal Using Dark Channel Prior

[2011 PAMI] Single Image Haze Removal Using Dark Channel Prior

7. [Denoising](#)

图像去噪也是图像处理中的一个经典问题, 在数码摄影中尤其重要。主要的方法有基于小波的方法和基于偏微分方程的方法。

[1992 SIAM] Image selective smoothing and edge detection by nonlinear diffusion. II

[1992 SIAM] Image selective smoothing and edge detection by nonlinear diffusion

[1992] Nonlinear total variation based noise removal algorithms

[1994 SIAM] Signal and image restoration using shock filters and anisotropic diffusion

[1995 TIT] De-noising by soft-thresholding

[1998 TIP] Orientation diffusions

[2000 TIP] Adaptive wavelet thresholding for image denoising and compression

[2000 TIP] Fourth-order partial differential equations for noise removal

[2001] Denoising through wavelet shrinkage

[2002 TIP] The Curvelet Transform for Image Denoising

[2003 TIP] Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time

[2008 PAMI] Automatic Estimation and Removal of Noise from a Single Image

[2009 TIP] Is Denoising Dead

8. [Edge Detection](#)

边缘检测也是图像处理中的一个基本任务。传统的边缘检测方法有基于梯度算子，尤其是 Sobel 算子，以及经典的 Canny 边缘检测。到现在，Canny 边缘检测及其思想仍在广泛使用。关于 Canny 算法的具体细节可以在 Sonka 的书以及 canny 自己的论文中找到，网上也可以搜到。最快最直接的方法就是看 OpenCV 的源代码，非常好懂。在边缘检测方面，Berkeley 的大牛 J Malik 和他的学生在 2004 年的 PAMI 提出的方法效果非常好，当然也比较复杂。在复杂度要求不高的情况下，还是值得一试的。MIT 的 Bill Freeman 早期的代表作 Steerable Filter 在边缘检测方面效果也非常好，并且便于实现。这里给出了几篇比较好的文献，包括一篇最新的综述。边缘检测是图像处理和计算机视觉中任何方向都无法逃避的一个问题，这方面研究多深都不为过。

[1980] theory of edge detection

[1983 Canny Thesis] find edge

[1986 PAMI] A Computational Approach to Edge Detection

[1990 PAMI] Scale-space and edge detection using anisotropic diffusion

[1991 PAMI] The design and use of steerable filters

[1995 PR] Multiresolution edge detection techniques

[1996 TIP] Optimal edge detection in two-dimensional images

[1998 PAMI] Local Scale Control for Edge Detection and Blur Estimation

[2003 PAMI] Statistical edge detection_ learning and evaluating edge cues

[2004 IEEE] Edge Detection Revisited

[2004 PAMI] Design of steerable filters for feature detection using canny-like criteria

[2004 PAMI] Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues

[2011 IVC] Edge and line oriented contour detection State of the art

9. [Graph Cut](#)

基于图割的图像分割算法。在这方面没有研究，仅仅列出几篇引用比较高的文献。这里又见 J Malik，当然还有华人杰出学者 Jianbo Shi，他的主页非常搞笑，在醒目的位置标注 [Do not fly China Eastern Airlines ...](#) 看来是被坑过，而且坑的比较厉害。这个领域，俄罗斯人比较厉害。

[2000 PAMI] Normalized cuts and image segmentation

[2001 PAMI] Fast approximate energy minimization via graph cuts

[2004 PAMI] What energy functions can be minimized via graph cuts

10. [Hough Transform](#)

虽然霍夫变换可以扩展到广义霍夫变换，但最常用的还是检测圆和直线。这方面同样推荐看 OpenCV 的源代码，一目了然。Matas 在 2000 年提出的 PPHT 已经集成到 OpenCV 中去了。

[1986 CVGIU] A Survey of the Hough Transform

[1989] A Comparative study of Hough transform methods for circle finding

[1992 PAMI] Shapes recognition using the straight line Hough transform_ theory and generalization

[1997 PR] Extraction of line features in a noisy image

[2000 CVIU] Robust Detection of Lines Using the Progressive Probabilistic Hough Transform

11. [Image Interpolation](#)

图像插值，偶尔也用得上。一般来说，双三次也就够了

[2000 TMI] Interpolation revisited

12. [Image Matting](#)

也就是最近，我才知道这个词翻译成中文是抠图，比较难听，不知道是谁开始这么翻译的。没有研究，请看文章以及 [Richard Szeliski](#) 的相关章节。以色列美女 Levin 在这方面有两篇 PAMI。

[2008 Fnd] Image and Video Matting A Survey

[2008 PAMI] A Closed-Form Solution to Natural Image Matting

[2008 PAMI] Spectral Matting

13. [Image Modeling](#)

图像的统计模型。这方面有一本专门的著作 Natural Image Statistics

[1994] The statistics of natural images

[2003 JMIV] On Advances in Statistical Modeling of Natural Images

[2009 IJCV] Fields of Experts

[2009 PAMI] Modeling multiscale subbands of photographic images with fields of Gaussian scale mixtures

14. [Image Quality Assessment](#)

在图像质量评价方面，Bovik 是首屈一指的。这位老师也很有意思，作为编辑出版了很多书。他也是 IEEE 的 Fellow

[2004 TIP] Image quality assessment from error visibility to structural similarity

[2011 TIP] blind image quality assessment From Natural Scene Statistics to Perceptual Quality

15. [Image Registration](#)

图像配准最早的应用在医学图像上，在图像融合之前需要对图像进行配准。在现在的计算机视觉中，配准也是一个需要理解的概念，比如跟踪，拼接等。在 KLT 中，也会涉及到配准。这里主要是综述文献。

[1992 MIA] Image matching as a diffusion process

[1992 PAMI] A Method for Registration of 3-D shapes

[1992] a survey of image registration techniques

[1998 MIA] A survey of medical image registration

[2003 IVC] Image registration methods a survey

[2003 TMI] Mutual-Information-Based Registration of Medical Survey

[2011 TIP] Hairis registration

16. [Image Retrieval](#)

图像检索曾经很热，在 2000 年之后似乎消停了一段时间。最近各种图像的不变性特征提出来之后，再加上互联网搜索的商业需求，这个方向似乎又要火起来了，尤其是在商业界，比如淘淘搜。这仍然是一个非常值得关注的方面。而且图像检索与目标识别具有相通之处，比如特征提取和特征降维。这方面的文章值得一读。在最后给出了两篇 Book chapter，其中一篇还是中文的。

[2000 PAMI] Content-based image retrieval at the end of the early years

[2000 TIP] PicToSeek Combining Color and Shape Invariant Features for Image Retrieval

[2002] Content-Based Image Retrieval Systems A Survey

[2008] Content-Based Image Retrieval-Literature Survey

[2010] Plant Image Retrieval Using Color, Shape and Texture Features

[2012 PAMI] A Multimedia Retrieval Framework Based on Semi-Supervised Ranking and Relevance Feedback

CBIR Chinese

fundament of cbir

17. [Image Segmentation](#)

图像分割, 非常基本但又非常难的一个问题。建议看 Sonka 和冈萨雷斯的书。这里给出几篇比较好的文章, 再次看到了 J Malik。他们给出了源代码和测试集, 有兴趣的话可以试试。

[2004 IJCV] Efficient Graph-Based Image Segmentation

[2008 CVIU] Image segmentation evaluation A survey of unsupervised methods

[2011 PAMI] Contour Detection and Hierarchical Image Segmentation

18. [Level Set](#)

大名鼎鼎的水平集, 解决了 Snake 固有的缺点。Level set 的两位提出者 Sethian 和 Osher 最后反目, 实在让人遗憾。个人以为, 这种方法除了迭代比较费时, 在真实场景中的表现让人生疑。不过, 2008 年 ECCV 上的 PWP 方法在结果上很吸引人。在重初始化方面, Chunming Li 给出了比较好的解决方案

[1995 PAMI] Shape modeling with front propagation_ a level set approach

[2001 JCP] Level Set Methods_ An Overview and Some Recent Results

[2005 CVIU] Geodesic active regions and level set methods for motion estimation and tracking

[2007 IJCV] A Review of Statistical Approaches to Level Set Segmentation

[2008 ECCV] Robust Real-Time Visual Tracking using Pixel-Wise Posteriors

[2010 TIP] Distance Regularized Level Set Evolution and its Application to Image Segmentation

19. [Pyramid](#)

其实小波变换就是一种金字塔分解算法, 而且具有无失真重构和非冗余的优点。Adelson 在 1983 年提出的 Pyramid 优点是比较简单, 实现起来比较方便。

[1983] The Laplacian Pyramid as a Compact Image Code

20. [Radon Transform](#)

Radon 变换也是一种很重要的变换，它构成了图像重建的基础。关于图像重建和 radon 变换，可以参考章毓晋老师的书，讲的比较清楚。

[1993 PAMI] Image representation via a finite Radon transform

[1993 TIP] The fast discrete radon transform I theory

[2007 IVC] Generalised finite radon transform for $N \times N$ images

21. [Scale Space](#)

尺度空间滤波在现代不变特征中是一个非常重要的概念，有人说 SIFT 的提出者 Lowe 是不变特征之父，而 Linderburg 是不变特征之母。虽然尺度空间滤波是 Witkin 最早提出的，但其理论体系的完善和应用还是 Linderburg 的功劳。其在 1998 年 IJCV 上的两篇文章值得一读，不管是特征提取方面还是边缘检测方面。

[1987] Scale-space filtering

[1990 PAMI] Scale-Space for Discrete Signals

[1994] Scale-space theory A basic tool for analysing structures at different scales

[1998 IJCV] Edge Detection and Ridge Detection with Automatic Scale Selection

[1998 IJCV] Feature Detection with Automatic Scale Selection

22. [Snake](#)

活动轮廓模型，改变了传统的图像分割的方法，用能量收缩的方法得到一个统计意义上的能量最小（最大）的边缘。

[1987 IJCV] Snakes Active Contour Models

[1996] deformable model in medical image A Survey

[1997 IJCV] geodesic active contour

[1998 TIP] Snakes, shapes, and gradient vector flow

[2000 PAMI] Geodesic active contours and level sets for the detection and tracking of moving objects

[2001 TIP] Active contours without edges

23. [Super Resolution](#)

超分辨率分析。对这个方向没有研究，简单列几篇文章。其中 Yang Jianchao 的那篇在 IEEE 上的下载率一直居高不下。

[2002] Example-Based Super-Resolution

[2009 ICCV] Super-Resolution from a Single Image

[2010 TIP] Image Super-Resolution Via Sparse Representation

24. [Thresholding](#)

阈值分割是一种简单有效的图像分割算法。这个 topic 在冈萨雷斯的书里面讲的比较多。这里列出 OTSU 的原始文章以及一篇不错的综述。

[1979 IEEE] OTSU A threshold selection method from gray-level histograms

[2001 JISE] A Fast Algorithm for Multilevel Thresholding

[2004 JEI] Survey over image thresholding techniques and quantitative performance evaluation

25. [Watershed](#)

分水岭算法是一种非常有效的图像分割算法，它克服了传统的阈值分割方法的缺点，尤其是 Marker-Controlled Watershed，值得关注。Watershed 在冈萨雷斯的书里面讲的比较详细。

[1991 PAMI] Watersheds in digital spaces an efficient algorithm based on immersion simulations

[2001]The Watershed Transform Definitions, Algorithms and Parallelizat on Strategies

五、 计算机视觉

这一章是计算机视觉部分，主要侧重在底层特征提取，视频分析，跟踪，目标检测和识别方面等方面。对于自己不太熟悉的领域比如摄像机标定和立体视觉，仅仅列出上 google 上引用次数比较多的文献。有一些刚刚出版的文章，个人非常喜欢，也列出来了。

1. [Active Appearance Models](#)

活动外观模型和活动轮廓模型基本思想来源 Snake，现在在人脸三维建模方面得到了很成功的应用，这里列出了三篇最早最经典的文章。对这个领域有兴趣的可以从这三篇文章开始入手。

[1998 ECCV] Active Appearance Models

[2001 PAMI] Active Appearance Models

2. [Active Shape Models](#)

[1995 CVIU] Active Shape Models—Their Training and Application

3. [Background modeling and subtraction](#)

背景建模一直是视频分析尤其是目标检测中的一项关键技术。虽然最近一直有一些新技术的产生，demo 效果也很好，比如基于 dynamical texture 的方法。但最经典的还是 Stauffer 等在 1999 年和 2000 年提出的 GMM 方法，他们最大的贡献在于不用 EM 去做高斯拟合，而是采用了一种迭代的算法，这样就不需要保存很多帧的数据，节省了 buffer。Zivkovic 在 2004 年的 ICPR 和 PAMI 上提出了动态确定高斯数目的方法，把混合高斯模型做到了极致。这种方法效果也很好，而且易于实现。在 OpenCV 中有现成的函数可以调用。在背景建模大家族里，无参数方法（2000 ECCV）和 Vibe 方法也值得关注。

[1997 PAMI] Pfister Real-Time Tracking of the Human Body

[1999 CVPR] Adaptive background mixture models for real-time tracking

[1999 ICCV] Wallflower Principles and Practice of Background Maintenance

[2000 ECCV] Non-parametric Model for Background Subtraction

[2000 PAMI] Learning Patterns of Activity Using Real-Time Tracking

[2002 PAMI] Background and foreground modeling using nonparametric

kernel density estimation for visual surveillance

[2004 ICPR] Improved adaptive Gaussian mixture model for background subtraction

[2004 PAMI] Recursive unsupervised learning of finite mixture models

[2006 PRL] Efficient adaptive density estimation per image pixel for the task of background subtraction

[2011 TIP] ViBe A Universal Background Subtraction Algorithm for Video Sequences

4. [Bag of Words](#)

词袋，在这方面暂时没有什么研究。列出三篇引用率很高的文章，以后逐步解剖之。

[2003 ICCV] Video Google A Text Retrieval Approach to Object Matching in Videos

[2004 ECCV] Visual Categorization with Bags of Keypoints

[2006 CVPR] Beyond bags of features Spatial pyramid matching for recognizing natural scene categories

5. [BRIEF](#)

BRIEF 是 Binary Robust Independent Elementary Features 的简称，是近年来比较受关注的特征描述的方法。ORB 也是基于 BRIEF 的。

[2010 ECCV] BRIEF Binary Robust Independent Elementary Features

[2011 ICCV] ORB an efficient alternative to SIFT or SURF

[2012 PAMI] BRIEF Computing a Local Binary Descriptor Very Fast

6. [Camera Calibration and Stereo Vision](#)

非常不熟悉的领域。仅仅列出了十来篇重要的文献，供以后学习。

[1979 Marr] A Computational Theory of Human Stereo Vision

[1985] Computational vision and regularization theory

[1987 IEEE] A versatile camera calibration technique for

high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses

[1987] Probabilistic Solution of Ill-Posed Problems in Computational Vision

[1988 PIEEE] Ill-Posed Problems in Early Vision

[1989 IJCV] Kalman Filter-based Algorithms for Estimating Depth from Image Sequences

[1990 IJCV] Relative Orientation

[1990 IJCV] Using vanishing points for camera calibration

[1992 ECCV] Camera self-calibration Theory and experiments

[1992 IJCV] A theory of self-calibration of a moving camera

[1992 PAMI] Camera calibration with distortion models and accuracy evaluation

[1994 IJCV] The Fundamental Matrix Theory, Algorithms, and Stability Analysis

[1994 PAMI] a stereo matching algorithm with an adaptive window theory and experiment

[1999 ICCV] Flexible camera calibration by viewing a plane from unknown orientations

[1999 IWAR] Marker tracking and hmd calibration for a video-based augmented reality conferencing system

[2000 PAMI] A flexible new technique for camera calibration

7. [Color and Histogram Feature](#)

这里面主要来源于图像检索，早期的图像检测基本基于全局的特征，其中最显著的就是颜色特征。这一部分可以和前面的 Color 知识放在一起的。

[1995 SPIE] Similarity of color images

[1996 PR] IMAGE RETRIEVAL USING COLOR AND SHAPE

[1996] comparing images using color coherence vectors

[1997] Image Indexing Using Color Correlograms

[2001 TIP] An Efficient Color Representation for Image Retrieval

[2009 CVIU] Performance evaluation of local colour invariants

8. [Deformable Part Model](#)

大红大热的 DPM，在 OpenCV 中有一个专门的 topic 讲 DPM 和 latent svm

[2008 CVPR] A Discriminatively Trained, Multiscale, Deformable Part Model

[2010 CVPR] Cascade Object Detection with Deformable Part Models

[2010 PAMI] Object Detection with Discriminatively Trained Part-Based Models

9. [Distance Transformations](#)

距离变换，在 OpenCV 中也有实现。用来在二值图像中寻找种子点非常方便。

[1986 CVGIP] Distance Transformations in Digital Images

[2008 ACM] 2D Euclidean Distance Transform Algorithms A Comparative Survey

10. [Face Detection](#)

最成熟最有名的当属 Haar+Adaboost

[1998 PAMI] Neural Network-Based Face Detection

[2002 PAMI] Detecting faces in images a survey

[2002 PAMI] Face Detection in Color Images

[2004 IJCV] Robust Real-Time Face Detection

11. [Face Recognition](#)

不熟悉，简单罗列之。

[1991] Face Recognition Using Eigenfaces

[2000 PAMI] Automatic Analysis of Facial Expressions The State of the Art

[2000] Face Recognition A Literature Survey

[2006 PR] Face recognition from a single image per person A survey

[2009 PAMI] Robust Face Recognition via Sparse Representation

12. [FAST](#)

用机器学习的方法来提取角点，号称很快很好。

[2006 ECCV] Machine learning for high-speed corner detection

[2010 PAMI] Faster and Better A Machine Learning Approach to Corner Detection

13. [Feature Extraction](#)

这里的特征主要都是各种不变性特征，SIFT, Harris, MSER 等也属于这一类。把它们单独列出来是因为这些方法更流行一点。关于不变性特征，王永明与王贵锦合著的《图像局部不变性特征与描述》写的还不错。Mikolajczyk 在 2005 年的 PAMI 上的文章以及 2007 年的综述是不错的学习材料。

[1989 PAMI] On the detection of dominant points on digital curves

[1997 IJCV] SUSAN—A New Approach to Low Level Image Processing

[2004 IJCV] Matching Widely Separated Views Based on Affine Invariant Regions

[2004 IJCV] Scale & Affine Invariant Interest Point Detectors

[2005 PAMI] A performance evaluation of local descriptors

[2006 IJCV] A Comparison of Affine Region Detectors

[2007 FAT] Local Invariant Feature Detectors - A Survey

[2011 IJCV] Evaluation of Interest Point Detectors and Feature Descriptors

14. [Feature Matching](#)

Fua 课题组在今年 PAMI 上的一篇文章，感觉还不错

[2012 PAMI] LDAHash Improved Matching with Smaller Descriptors

15. [Harris](#)

虽然过去了很多年，Harris 角点检测仍然广泛使用，而且基于它有很多变形。如果仔细看了这种方法，从直观也可以感觉到这是一种很稳健的方法。

[1988 Harris] A combined corner and edge detector

16. [Histograms of Oriented Gradients](#)

HoG 方法也在 OpenCV 中实现了：HOGDescriptor。

[2005 CVPR] Histograms of Oriented Gradients for Human Detection
NavneetDalalThesis.pdf

17. [Image Distance](#)

[1993 PAMI] Comparing Images Using the Hausdorff Distance

18. [Image Stitching](#)

图像拼接，另一个相关的词是 Panoramic。在 Computer Vision: Algorithms and Applications 一书中，有专门一章是讨论这个问题。这里的两面文章一篇是综述，一篇是这方面很经典的文章。

[2006 Fnd] Image Alignment and Stitching A Tutorial

[2007 IJCV] Automatic Panoramic Image Stitching using Invariant Features

19. [KLT](#)

KLT 跟踪算法，基于 Lucas-Kanade 提出的配准算法。除了三篇很经典的文章，最后一篇给出了 OpenCV 实现 KLT 的细节。

[1981] An Iterative Image Registration Technique with an Application to Stereo Vision full version

[1994 CVPR] Good Features to Track

[2004 IJCV] Lucas-Kanade 20 Years On A Unifying Framework

20. [Local Binary Pattern](#)

LBP。OpenCV 的 Cascade 分类器也支持 LBP，用来取代 Haar 特征。

[2002 PAMI] Multiresolution gray-scale and rotation Invariant Texture Classification with Local Binary Patterns

[2004 ECCV] Face Recognition with Local Binary Patterns

[2006 PAMI] Face Description with Local Binary Patterns

[2011 TIP] Rotation-Invariant Image and Video Description With Local Binary Pattern Features

21. [Low-Level Vision](#)

关于 Low level vision 的两篇很不错的文章

[1998 TIP] A general framework for low level vision

[2000 IJCV] Learning Low-Level Vision

22. [Mean Shift](#)

均值漂移算法，在跟踪中非常流行的方法。Comaniciu 在这个方面做出了重要的贡献。最后三篇，一篇是 CVIU 上的 top download 文章，一篇是最新的 PAMI 上关于 Mean Shift 的文章，一篇是 OpenCV 实现的文章。

[1995 PAMI] Mean shift, mode seeking, and clustering

[2002 PAMI] Mean shift a robust approach toward feature space analysis

[2003 CVPR] Mean-shift blob tracking through scale space

[2009 CVIU] Object tracking using SIFT features and mean shift

[2012 PAMI] Mean Shift Trackers with Cross-Bin Metrics

OpenCV Computer Vision Face Tracking For Use in a Perceptual User Interface

23. [MSER](#)

这篇文章发表在 2002 年的 BMVC 上，后来直接录用到 2004 年的 IVC 上，内容差不多。MSER 在 Sonka 的书里面也有提到。

[2002 BMVC] Robust Wide Baseline Stereo from Maximally Stable Extremal Regions

[2003] MSER Author Presentation

[2004 IVC] Robust wide-baseline stereo from maximally stable extremal regions

[2011 PAMI] Are MSER Features Really Interesting

24. [Object Detection](#)

首先要说的是第一篇文章的作者，Kah-Kay Sung。他是 MIT 的博士，后来到新加坡国立任教，极具潜力的一个老师。不幸的是，他和他的妻子都在 2000 年的新加坡空难中遇难，让人唏嘘不已。

http://en.wikipedia.org/wiki/Singapore_Airlines_Flight_006

最后一篇文章也是 Fua 课题组的，作者给出的 demo 效果相当好。

[1998 PAMI] Example-based learning for view-based human face detection

[2003 IJCV] Learning the Statistics of People in Images and Video

[2011 PAMI] Learning to Detect a Salient Object

[2012 PAMI] A Real-Time Deformable Detector

25. [Object Tracking](#)

跟踪也是计算机视觉中的经典问题。粒子滤波，卡尔曼滤波，KLT, mean shift, 光流都跟它有关系。这里列出的是传统意义上的跟踪，尤其值得一看的是 2008 年的 Survey 和 2003 年的 Kernel based tracking。

[2003 PAMI] Kernel-based object tracking

[2007 PAMI] Tracking People by Learning Their Appearance

[2008 ACM] Object Tracking A Survey

[2008 PAMI] Segmentation and Tracking of Multiple Humans in Crowded Environments

[2011 PAMI] Hough Forests for Object Detection, Tracking, and Action Recognition

[2011 PAMI] Robust Object Tracking with Online Multiple Instance Learning

[2012 IJCV] PWP3D Real-Time Segmentation and Tracking of 3D Objects

26. [OCR](#)

一个非常成熟的领域，已经很好的商业化了。

[1992 IEEE] Historical review of OCR research and development

Video OCR A Survey and Practitioner's Guide

27. [Optical Flow](#)

光流法，视频分析所必需掌握的一种算法。

[1981 AI] Determine Optical Flow

[1994 IJCV] Performance of optical flow techniques

[1995 ACM] The Computation of Optical Flow

[2004 TR] Tutorial Computing 2D and 3D Optical Flow

[2005 BOOK] Optical Flow Estimation

[2008 ECCV] Learning Optical Flow

[2011 IJCV] A Database and Evaluation Methodology for Optical Flow

28. [Particle Filter](#)

粒子滤波，主要给出的是综述以及 1998 IJCV 上的关于粒子滤波发展早期的经典文章。

[1998 IJCV] CONDENSATION—Conditional Density Propagation for Visual Tracking

[2002 TSP] A tutorial on particle filters for online nonlinear non-Gaussian Bayesian tracking

[2002 TSP] Particle filters for positioning, navigation, and tracking

[2003 SPM] particle filter

29. [Pedestrian and Human detection](#)

仍然是综述类，关于行人和人体的运动检测和动作识别。

[1999 CVIU] Visual analysis of human movement_ A survey

[2001 CVIU] A Survey of Computer Vision-Based Human Motion Capture

[2005 TIP] Image change detection algorithms a systematic survey

[2006 CVIU] a survey of advances in vision based human motion capture

[2007 CVIU] Vision-based human motion analysis An overview

[2007 IJCV] Pedestrian Detection via Periodic Motion Analysis

[2007 PR] A survey of skin-color modeling and detection methods

[2010 IVC] A survey on vision-based human action recognition

[2012 PAMI] Pedestrian Detection An Evaluation of the State of the Art

30. [Scene Classification](#)

当相机越来越傻瓜化的时候，自动场景识别就非常重要。这是比拼谁家的Auto功能做的比较好的时候了。

[2001 IJCV] Modeling the Shape of the Scene A Holistic Representation of the Spatial Envelope

[2001 PAMI] Visual Word Ambiguity

[2007 PAMI] A Thousand Words in a Scene

[2010 PAMI] Evaluating Color Descriptors for Object and Scene Recognition

[2011 PAMI] CENTRIST A Visual Descriptor for Scene Categorization

31. [Shadow Detection](#)

[2003 PAMI] Detecting moving shadows-- algorithms and evaluation

32. [Shape](#)

关于形状，主要是两个方面：形状表示和形状识别。形状表示主要是从边缘或者区域当中提取不变性特征，用来做检索或者识别。这方面 Sonka 的书讲的比较系统。2008 年的那篇综述在这方面也讲的不错。至于形状识别，最牛的当属 J Malik 等提出的 Shape Context。

[1993 PR] IMPROVED MOMENT INVARIANTS FOR SHAPE DISCRIMINATION

[1993 PR] Pattern Recognition by Affine Moment Invariants

[1996 PR] IMAGE RETRIEVAL USING COLOR AND SHAPE

[2001 SMI] Shape matching similarity measures and algorithms

[2002 PAMI] Shape matching and object recognition using shape contexts

[2004 PR] Review of shape representation and description techniques

[2006 PAMI] Integral Invariants for Shape Matching

[2008] A Survey of Shape Feature Extraction Techniques

33. [SIFT](#)

关于 SIFT，实在不需要介绍太多，一万多次的引用已经说明问题了。SURF 和 PCA-SIFT 也是属于这个系列。后面列出了几篇跟 SIFT 有关的问题。

[1999 ICCV] Object recognition from local scale-invariant features

[2000 IJCV] Evaluation of Interest Point Detectors

[2003 CVIU] Speeded-Up Robust Features (SURF)

[2004 CVPR] PCA-SIFT A More Distinctive Representation for Local Image Descriptors

[2004 IJCV] Distinctive Image Features from Scale-Invariant Keypoints

[2010 IJCV] Improving Bag-of-Features for Large Scale Image Search

[2011 PAMI] SIFTflow Dense Correspondence across Scenes and its Applications

34. [SLAM](#)

Simultaneous Localization and Mapping, 同步定位与建图。

SLAM 问题可以描述为：机器人在未知环境中从一个未知位置开始移动, 在移动过程中根据位置估计和地图进行自身定位, 同时在自身定位的基础上建造增量式地图, 实现机器人的自主定位和导航。

[2002 PAMI] Simultaneous Localization and Map-Building Using Active Vision

[2007 PAMI] MonoSLAM Real-Time Single Camera SLAM

35. [Texture Feature](#)

纹理特征也是物体识别和检索的一个重要特征集。

[1973] Textural features for image classification

[1979] Statistical and structural approaches to texture

[1996 PAMI] Texture features for browsing and retrieval of image data

[2002 PR] Brief review of invariant texture analysis methods

[2012 TIP] Color Local Texture Features for Color Face Recognition

36. [TLD](#)

Kadal 创立了 TLD, 跟踪学习检测同步进行, 达到稳健跟踪的目的。他的两个导师也是大名鼎鼎, 一个是发明 MSER 的 Matas, 一个是 Mikolajczyk。他还创立了一个公司 [TLD Vision s.r.o.](#) 这里给出了他的系列文章, 最后一篇是刚出来的 PAMI。

[2009] Online learning of robust object detectors during unstable tracking

[2010 CVPR] P-N Learning Bootstrapping Binary Classifiers by Structural Constraints

[2010 ICIP] FACE-TLD TRACKING-LEARNING-DETECTION APPLIED TO FACES

[2012 PAMI] Tracking-Learning-Detection

37. [Video Surveillance](#)

前两篇是两个很有名的视频监控系统，里面包含了很丰富的信息量，比如 CMU 的那个系统里面的背景建模算法也是相当简单有效的。最后一篇是比较近的综述。

[2000 CMU TR] A System for Video Surveillance and Monitoring

[2000 PAMI] W4-- real-time surveillance of people and their activities

[2008 MVA] The evolution of video surveillance an overview

38. [Viola-Jones](#)

Haar+Adaboost 的弱弱联手，组成了最强大的利器。在 OpenCV 里面有它的实现，也可以选择用 LBP 来代替 Haar 特征。

[2001 CVPR] Rapid object detection using a boosted cascade of simple features

[2004 IJCV] Robust Real-time Face Detection

六、 结束语

历时一个多月，终于用业余时间把这些资料整理出来了，总算了却了一块心病，也不至于再看着一堆资料发愁了。以后可能会有些小修小补，但不会有太大的变化了。万里长征走完了第一步，剩下的就是理解和消化了。借新浪 [ishare](#) 共享出来，希望能够对你的科研也有一定的帮助。最后简单统计一下各个年份出现的频率。

文章总数： 372

2012 年： 10

2011 年： 20

2010 年： 20

2009 年： 14

2008 年： 18

2007 年： 13

2006 年： 14

2005 年： 9

2004 年： 24

2003 年： 22

2002 年： 21

2001 年： 21

2000 年： 23

1999 年： 10

1998 年： 22

1997 年： 8

1996 年： 9

1995 年： 9

1994 年： 7

1993 年： 5

1992 年： 11

1991 年： 5

1990 年： 6

1980-1989: 22

1960-1979: 9