****

**J I A N G S U U N I V E R S I T Y**

**本 科 毕 业 论 文**

基于ViBe算法的视频浓缩系统设计

**Design of Video Enrichment System Based on ViBe Algorithm**

学院名称： 计算机科学技术与通信工程

专业班级： 计算机1301

学生姓名： 金洪影

指导教师姓名： 詹永照

指导教师职称： 教授

2017年 6 月

基于ViBe算法的视频浓缩系统设计

专业班级：计算机1101 学生姓名：金洪影

指导教师：詹永照 职称：教授

摘要

# 第一章 绪论

## 1.1 研究背景及意义

近年来，人们的生活水平得到了迅速的提高，与此同时，社会也变得越来越复杂，社会中的信息量日益庞大，如何从如此繁重的信息中提取出有用的信息就成了我们所要面对的一个现实问题。随着社会中监控设备的普及，监控得到的视频信息也变得越来越多，在监控时实时分析庞大的视觉信息，然后保存其中重要的信息逐渐受到广大人民的重视 。在城市交通、安全监控的应用中，视频数据急速增长，视频信息已然成为一个海量数据资源。海量的视频数量给视频的存储和归档带了巨大压力，同时监控视频在时空上存在冗余，降低了了工作人员调阅视频的速度并加大了其难度。因此考虑通过视频浓缩技术对原始视频体积和长度进行压缩，从而达到减少视频存储压力，降低人工查阅的难度的目标。

监控系统的发展过程经历了三个发展发展阶段，分别是第一代的全模拟（VCR），第二代的部分数字化（DVR/NVR）和第三代的全数字化系统。虽然在这个过程中视频监控系统设备的性能得到了很大的提高，但是其还是存在一些缺点和限制，比如在一些比较敏感的场所，如银行、机场、交通要道等，从安全的角度出发，管理者需要对其进行全天候的监控，从而可以发现一些突发的异常状况，然后对其采取相应的措施。但是这也就相应地产生了一个问题，即监视者体力下降的同时会出现注意力分散、反应速度慢、身体疲劳等情况，进而影响工作效率。

背景差分（BS）技术是自动视频分析的关键技术之一，特别是在视频监控领域。 早期静态背景模型可能适用于在受限室内环境中分析短视频序列，但该模型对于大多数实际情况是无效的; 因此需要更复杂的模型。 此外，运动的检测往往只是理解现场的第一步。 例如，用于检测无人看管行李区、步态识别、人脸检测、人数统计、交通监控等风光背景时，检测到的运动可能会被表征和过滤。过去几年里，已经开发了各种BS方法（参考[24,9,4,2,3]），每个方法都有自己的特点，长处和弱点。评估允许识别这些特征，并有助于专注于剩余的问题。虽然文献重要，但是文献缺乏对近期BS方法的综合评价。一个原因可能是在自然视频序列生成定性高地面真相（GT）数据方面付出了巨大努力。因此，一些评估仅使用几个标记的框架或判断对象层面的性能，是相当容易的。然而，像素级的评估提供了更多的洞察力量和弱点。存在克服手动GT注释的几种技术。除了人力专家对分割结果的劳动密集和高度主观判断外，还开发出不依赖于GT数据[5,10]或自动生成它们的各种方法[12]。不幸的是，这种方法不适用于对视频监控中出现的挑战评估BS性能（参见第3节）。为了解决GT数据采集的问题，我们提出使用人工数据。为了应对“综合数据可能不会忠实地代表真实数据的全面数据”的问题[9]，我们使用典型的视频监控场景，高质量的3D模型和具有全局照明的现代光线跟踪技术用于逼真的图像合成。因此，我们能够生成高质量的像素级GT数据，并与别人分开评估挑战。基于引入的数据集，我们通过后处理来评估9种BS方法的性能。这包括一些多模式的方法，能够应对动态背景。在现有评估中几乎没有比较这种方法。此外，我们能够应对近期评估文献（例如阴影和噪音）几乎没有涵盖的挑战。BS的主要任务是将输入帧与背景模型进行比较。请注意，此处忽略了诸如图像配准，颜色转换等输入的任何预处理。该模型描述了场景的背景区域，并且通常由颜色信息等特征的分布来表示。前景检测的过程根据输入帧和背景模型的相似度确定图像的哪些区域属于前景类。

更佳的解决方案是提供一种从单个帧初始化背景模型的技术，一种视频序列通用的BS算法—ViBe，油然而生。

## 1.2 国内外研究现状

目前国内外在视频浓缩方面取得了大量的科研成果。中科院和中科奥森合作的视频浓缩快览产品能针对长时间视频在时间和空间上进行高效的浓缩，能够保证不丢失视频中的重要运动目标信息。可以快速捕捉视频内容和掌握事件重要信息。视频的浓缩速度可以达到8倍以上。高度压缩，场景标准[a]。王海军、葛红娟、张圣燕针对时间、空间的冗余，提出视频拼接技术[b]，其思想是使用运动目标检测及跟踪技术，将不同空间、时间的视频帧乃至于其中的对象衔接在一起。以色列耶路撒冷希伯来大学选用基于对象的视频摘要技术，经过切割视频中得运动物体，再对其实行时间平移后进行重新组合的方法，研制开发了Brief Cam[c],主要用于海量监控视频内容的浓缩和检索。利用该技术，可以将5个小时的视频浓缩成5分钟。

背景差分技术是背景建模的关键技术。Herrero和Bescos提出了使用人工组合的前景和背景对象的广泛数据集对BS技术的系统评估。然而，这个数据集不包括像场景照明和阴影这样的挑战。

ViBe算法是视频序列的一种通用的背景差分算法，完成视频背景模型的建立，是本次系统设计的核心所在。至今为止，已经提出了许多背景差分技术，具有许多模型和分割策略。一些算法集中在理想的背景差分技术可以或应该实现的具体要求。根据[7]，背景差分技术必须适应逐渐或快速的照明变化（改变日，云等），运动变化（摄像机 振荡），高频背景物体（例如树木离开或分支）以及背景几何体（例如停放的汽车）的变化。一些应用需要将背景差分算法嵌入到相机中，使计算负载成为主要关注的问题。对于室外场景的监视，对噪声的鲁棒性和对照明变化的适应性也是至关重要的。

## 1.3 本论文研究的内容及主要工作

此次系统设计借助OpenCV提供的视频图像处理库，实现改进的ViBe算法，逐帧对视频进行背景建模，通过背景差法，判断每一帧是否含有感兴趣的前景对象，如果不存在，则将该视频帧略去，只保留含有前景对象的视频帧，最后将保留下来的视频帧整合成新的视频文件，达到视频浓缩的效果[d]。具体分为以下几个模块：（1）读取视频文件；（2）对视频序列进行背景建模；（3）对视频帧进行划分；（4）重新整合视频文件；（5）在基础视频上再浓缩。

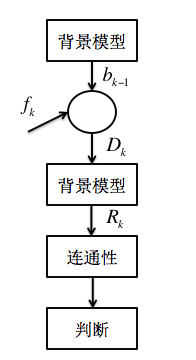
本文提出了在原有的ViBe算法上进行改进，不同的距离函数和阈值准则，使用不同大小的形态学方法，抑制updating mask中某些像素的传播，检测闪烁像素点，当相机有抖动时，增加更新因子。在改进的ViBe上进行背景建模并完成初步浓缩视频后，采用感知哈希算法来比较每一帧图像的相似度。形成更加浓缩的视频。

## 1.4 论文的内容安排

# 第二章 背景差分相关的技术研究现状

## 2.1 背景差分核心理论

普遍的定义是将静态背景帧与视频场景的当前帧逐像素进行比较。这是背景差分的基本原理，可以将其形成为构建背景模型的技术，并将该模型与当前帧进行比较，以便检测出显着差异的区域。 因此，背景减法算法的目的是将移动对象（以下称为前景）与场景的静态或慢移动部分（称为背景）区分开。由背景差分技术处理的问题涉及将观察到的图像与不包含任何感兴趣对象的估计图像进行比较;这被称为背景模型（或背景图像）[3]。 该比较过程称为前景检测，将观察到的图像划分为覆盖整个图像的两个补充的像素集合，（1）包含感兴趣对象的前景；（2）背景及其互补集。 背景建模原理图如图所示：



## 2.2 相关技术

### 2.2.1 PCA模型

大多数技术独立地对每个像素进行操作。这些技术完全将后处理算法转化为对其结果添加某种形式的空间一致性的任务。由于扰动通常会影响单个像素，因此会导致局部错误分类。相比之下，Seiki等人描述的方法[10]，基于这样的假设，即相邻的背景像素块应当随时间而变化。虽然这个假设大部分时间都是有效的，特别是对于属于相同背景对象的像素而言，对于位于多个背景对象边界的相邻像素来说，这是一个问题。尽管这种不便，像素被聚合成块，并且每个N×N块被处理为一个N的平方分量矢量。然后随着时间的推移收集一些样本，并用于训练每个块的主成分分析（PCA）模型。如果其观察到的图像模式接近于使用8个相邻块的PCA投影系数的重建，则将新视频帧的块分类为背景。这种技术也在[11]中描述，但是它缺少一种更新机制来适应块模型随时间的变化。在[12]中，作者侧重于PCA重建误差。虽然PCA模型也训练有时间样本，结果模型占整个图像。使用在其PCA系数的图像空间中的当前图像和反投影之间的简单图像差分阈值将单个像素分类为背景或前景。对于其他基于PCA的方法，不会描述初始化过程和更新机制。

### 2.2.2 高斯混合模型

多年来，已经提出了越来越复杂的像素级算法。其中，目前最流行的是高斯混合模型（GMM）[40]，[41]。首先在[40]中提出，该模型包括通过高斯加权混合来建模每个像素随时间观察的值的分布。这种背景像素模型能够应对许多实际情况的多模式性质，并且当遇到重复的背景运动（如树叶或树枝）时，可以获得良好的效果。自引入以来，该模型在计算机视觉社区[4]，[7]，[11]，[42] - [44]中获得了广泛的普及，作者继续重访，这仍然引起了很大的兴趣该方法并提出增强算法[45] - [50]。在[51]中，提出了一种粒子群优化方法来自动确定GMM算法的参数。 [52]的作者将GMM模型与基于颜色直方图和纹理信息的基于区域的算法相结合。在他们的实验中，作者的方法优于原始的GMM算法。然而，作者的技术具有相当大的计算成本，因为它们只能使用Intel Xeon 5150处理器来处理每秒640×480像素的七帧。GMM的更新权值大小对背景更新有着很大的影响，如图所示：





图.采用不同的权值对背景更新影响的比较

根据实验选择合适的更新权值。GMM算法的缺点在于其强烈的假设，即背景比前景更为可见，并且其方差显着降低。 这些都不适用于每个时间窗口。 此外，如果在背景中存在高频和低频变化，则其灵敏度不能被精确调整，并且模型可以适应目标本身或者错过某些高速目标的检测，如[53]中详述的。 此外，模型参数（特别是方差）的估计在现实世界的嘈杂环境中可能会成为问题。 这通常使一个没有别的选择，而不是在硬件实现中使用固定的方差。 最后，应该指出的是，高斯模型的统计相关性是有争议的，因为一些作者声称自然图像表现出非高斯统计[54]。

### 2.2.3 码本模型

在码本算法[61]，[62]中，每个像素由码本表示，码本是长图像序列的背景模型的压缩形式。 每个码本由包括由创新的色彩失真度量变换的颜色的码字组成。 [63]提出了一种结合每个像素的空间和时间背景的改进码本。 相信代码本可以在有限的内存中长时间捕获背景运动。 因此，代码簿从通常长的训练序列中学习，并且在[62]中描述了码本更新机制，允许算法随着训练阶段结束时随着照明条件演变。 然而，应该注意到，建议的码本更新机制不允许创建新的码字，如果在后台发生永久的结构变化（例如，在城市户外场景中新释放的停车位的情况下），这可能是有问题的。

[64]，[65]的作者不是选择一种特定形式的背景密度模型，而是使用“共识”的概念。它们保留每个像素的给定数量的最后观察到的背景值的缓存，并将新值分类为背景，如果它匹配存储在像素模型中的大多数值。人们可能会期望这样一种方法可以避免与任意假定的密度模型的偏差有关的问题，但是由于像先进先出的更新策略来替换像素模型的值，所以它们也容易出现以前讨论的问题例如，除非存储大量的像素样本，否则背景中的慢和快速运动的问题。作者指出，20个样本的缓存是该方法有用的最低要求，但是他们也注意到超过60个样本的缓存没有显着的进一步改进。因此，其算法的训练周期必须至少包含20帧。最后，为了应对在背景中出现或衰落的照明变化和对象，在共享算法中添加了两个附加机制（一个在像素级别，一个在blob级别），以处理整个对象。

## 2.3 本课题面临的问题与挑战

## 2.4 小结

# 第三章 ViBe算法的改进及新方法模型

背景差分技术必须处理至少三个考虑因素，以便在实际应用中取得成功：（1）什么是模型，它的运作方式如何？（2）模型如何初始化？（3）模型随时间更新？这些问题的答案在本节的三个小节中给出。大多数论文描述了内在模型和更新机制。只有少数文章讨论初始化，这在预期快速响应时是至关重要的，如在数码相机中的情况。此外，模型和更新机制之间往往缺乏一致性。例如，一些技术将像素p的当前值与具有给定公差T的模型b的当前值进行比较。他们认为如果p和b之间的绝对差低于T，则存在很好的匹配。为了随着时间的推移自适应，相对于p的统计方差来调整T。但是统计方差是通过时间平均估计的。因此，调整速度取决于采集帧率和背景像素数。这在某些情况下是不合适的，如在其帧率由可用带宽确定的远程IP摄像机的情况下。

在总结了其他背景建模算法的优点后，提出了“ViBe”（“VIsual Background Extractor”）优化算法的背景差分技术。

## 3.1 ViBe算法理论

在某种程度上，对于给定的颜色空间，对于每个背景像素的确定概率密度函数（pdf）或至少统计参数的确定（诸如均值或方差）是无法确定的。注意，使用高斯模型，没有区别，因为平均值和方差的知识足以确定pdf。虽然背景扣除和大多数主流技术的经典方法依赖于pdf或统计参数，但是如果不是简单地忽略其统计意义的问题，很少讨论。事实上，只要实现相关背景细分的目标，就不必计算pdf。另一种方法是考虑随着时间推移增加统计学意义，一种方法就是建立具有实际观察像素值的模型。潜在的假设是，从随机的角度来看，这是更有意义的，因为已经观察到的值应该比没有遇到的值再次被观察的可能性更高。

像[65]的作者一样，我们不选择使用特定的pdf格式，因为与假定的pdf模型的偏差是普遍存在的。此外，对pdf的评估是一个全球性的过程，并且PDF的形状对异常值敏感。此外，pdf的估计提出了关于待考虑样本数量的不明显问题;选择代表数量的样本的问题是所有估计过程的固有的。

  如果我们看到背景减法的问题是一个分类问题，我们想在所选颜色空间中对其邻近的邻域分类一个新的像素值，以避免任何异常值的影响。这促使我们用一组样本而不是显式像素模型对每个背景像素进行建模。因此，不执行背景像素的pdf的估计，因此将像素的当前值与样本集合内的最接近的样本进行比较。与现有算法相比，这是一个重要的区别，特别是与基于共识的技术相比。将一个新值与背景样本进行比较，并应接近某些样本值，而不是所有值的大部分。基本思想是，估计具有少量接近值的背景像素的统计分布比大量样本更可靠。这有点类似于忽略pdf的四肢，或者仅通过阈值来考虑基础pdf的中心部分。另一方面，如果人们相信模型的值，那么仔细选择背景像素样本就是至关重要的。因此，背景中像素的分类需要保守，因为只有背景像素才能填充背景模型。

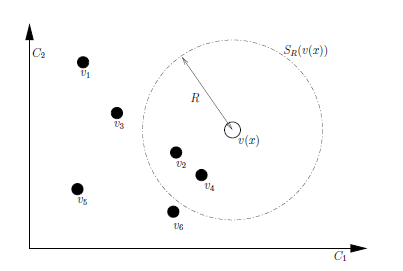


图1.像素值与二维欧几里德色彩空间（C1，C2）中的一组样本的比较。 为了对v（x）进行分类，我们计算以v（x）为中心的半径R的球体相交的M（x）的样本数。

形式上，让我们用v（x）表示由位于图像中x处的像素取得的给定欧几里得颜色空间中的值，以及vi表示具有索引i的背景样本值。 每个背景像素x由先前帧中采集的N个背景样本值的集合建模。 现在，我们忽略时间的概念; 这将在后面讨论。

M(x) = {v1, v2, . . . , vN} (1)

为了根据其对应的模型M（x）对像素值v（x）进行分类，我们通过定义以v（x）为中心的半径R的球体SR（v（x））将其与采样集合中的最接近的值进行比较）。 如果该球体的集合交集与模型样本集合M（x）的基数表示为大于或等于给定阈值 #min，则将像素值v（x）分类为背景。 更正式地，我们比较 #min到

#{SR（v（x））\ {v1，v2，。。。 ，vN}}（2）

根据等式2，像素值v（x）的分类涉及v（x）和模型样本之间的N个距离的计算，以及与阈值欧几里得距离R的N比较的计算。这个过程如图1所示。请注意，由于我们只想找几个匹配，一旦发现#min匹配，就可以停止像素的分割过程。

可以很容易地看出，我们的模型的精度由两个参数决定：球体半径R和最小基数 #min。 实验表明，20（单色图像）的独特半径R和2的基数是适当的。 在背景减法期间不需要调整这些参数，也不需要为图像中的不同像素位置更改它们。 注意，由于采样数N和 #min被选择为固定，并且由于它们影响相同的决定，所以可以使用以下比例来调整模型的灵敏度

 （3），

但是在所有的比较测试中，我们保持这些值不变。

### 3.1.1单帧初始化背景

ViBe提供一种从单个帧初始化背景模型的技术。给出这样的技术，对突然的照明变化的响应是直接的：现有的背景模型被丢弃并且新的模型被瞬时地初始化。此外，能够在序列的第二帧早期提供可靠的前景分割对于视频监视中的短序列或嵌入运动检测算法的装置具有明显的益处。由于在单个帧中没有时间信息，我们使用与[66]的作者相同的假设，即相邻像素具有类似的时间分布。这证明了我们使用在每个像素的空间邻域中找到的值来填充像素模型的事实。更准确地说，我们用第一帧中随机抽取的值填充它们。需要选择邻域的大小，使其足够大以构成足够数量的不同样本，同时记住不同位置处的值之间的统计学相关性随着邻域尺寸的增加而减小。从我们的实验中，在每个像素的8连接的邻域中随机选择样本已被证明对于640×480像素的图像是令人满意的。



根据公式（4），我们假设t = 0指示第一帧，并且NG（x）是像素位置x的空间邻域，因此根据均匀定律随机选择位置y。 注意，给定的v0（y）可以被选择几次（例如，如果邻域的大小小于M0（x）的基数）或者根本不被选择。 这个策略已经证明是成功的。 唯一的缺点是在第一帧中存在移动物体将引入通常称为鬼影。 根据[24]，鬼是“一组连接的点，被检测为运动但不对应于任何真正的移动物体”。 在这种特殊情况下，鬼是由具有来自移动物体的样本的像素模型的不幸初始化引起的。 在随后的帧中，对象移动并揭开真实的背景，这将通过常规模型更新过程逐渐学习，使幽灵随着时间的推移逐渐消失。 幸运的是，如第IV-C节所示，我们的更新过程确保在存在幽灵的情况下快速恢复模型，并将真正的移动对象缓慢并入背景模型。

### 3.1.2更新背景模型

保守的更新策略从不包含属于背景模型中的前景区域的样本。实际上，只有当像素样本被分类为背景样本时，才能将其包含在背景模型中。这样的政策看起来似乎是一个明显的选择，实际上它保证了对移动物体的清晰检测，因为它们不具有与背景相似的颜色。不幸的是，它也导致死锁情况和永恒的鬼影：背景样本被错误地分类为前景，导致无限期地保持背景像素模型不被更新并且可能导致永久错误分类。后面还会介绍在更新策略中做出的一些改进。

更新方法包括三个重要组成部分：（1）无记忆更新策略，确保存储在背景像素模型中的样本有平滑衰减寿命；（2）随机时间重采样，延长背景像素模型覆盖的背景像素；（3）在空间上传播背景像素样本，是确保空间一致性并允许对由前景屏蔽的背景像素模型的适应机制。在以下三个小节中，将阐述这些组件以及应用它们的原因。

1. 无记忆更新策略：为了在场景背景下处理广泛的事件，Wang[65]提出在像素模型中包含大量样本。但是如前所述，对于高帧率来说，这仍然是不适合的。其他作者[53]，[58]结合了两个时间子模型来处理快速和慢速修改。这种方法被证明是有效的。然而，它增加了参数化问题，因此有必要确定更多数量的参数值来实现。

这里提出一个方法，该方法通过允许几个旧样本保留在像素模型中来提高估计的时间相关性。这种方法与保守的更新策略相结合，因此前景值不应该包含在模型中。

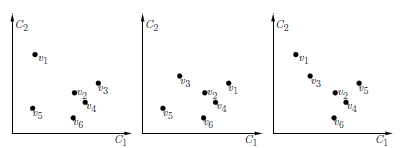


图.更新后的3种不同模型

由图可以看出，在大小N = 6的像素模型的更新了6个可能结果中的3个。假设像素值占据与第一个图相同的颜色空间，并且更新模型。 该图显示了更新后的3种可能的模型。 选择一个特定模型的决策过程是随机的（具有相同的概率）。显然该技术是简单但有效的，代替了系统地从像素模型中去除最旧的样本，根据均匀的概率密度函数选择随机丢弃的样本。

这种随机策略与先前更换较旧值的观念相矛盾，对于保守的更新策略来说，这并不适用。 保守的更新策略对于流程的稳定性也是必需的。 实际上，随机更新策略产生非确定性背景减法算法。 只有保守的更新政策才能确保模型不会随着时间而偏离。 尽管如此，在经验中，在不同时间由我们的背景减法算法处理的相同序列，结果之间可能存在微小差异。

在数学上，在像素模型更新之后在时间存在的模型中存在的样本的概率由给出。 假设时间连续性和选择过程中不存在存储器，我们可以在任何进一步的时间下得出类似的概率，表示为。这个概率表达式为



也可表示为：



该表达式表明，模型的任何样本值的预期剩余寿命呈指数衰减。 假设样本在时间t之前被包含在模型中的间隔（t，t + dt）似乎保留的概率与t无关。 换句话说，过去对未来没有影响。 这种称为无记忆性质的性质已知适用于指数密度（参见[68]）。 这是一个非凡的，并且据我们所知，背景扣除领域的独特性质。 它完全释放了我们定义一个将样本保持在像素历史中的时间段，并且在某种程度上允许更新机制适应任意帧率。

2）时间重新采样：我们已经展示了如何使用随机替换策略来允许我们的像素模型覆盖有限数量样本的大的（理论上无限的）时间窗口。 为了进一步扩展由固定大小的像素模型覆盖的时间窗口的大小，我们采用随机时间子采样。 这个想法是，在许多实际情况下，不必更新每个新帧的每个背景像素模型。 通过使后台更新频繁，我们人为地延长了后台样本的预期寿命。 但是在存在周期性或伪周期背景运动的情况下，使用固定的子采样间隔可能会阻止背景模型适当地适应这些运动。 这促使我们使用随机子采样策略。 实际上，当像素值已被归类为属于背景时，随机过程确定该值是否用于更新相应的像素模型。

在所有测试中，采用时间重采样因子，默认表示为：16，背景像素值在16个被选中更新其像素模型时有1次机会。但是，您可能希望调整此参数来调整像素模型覆盖的时间窗口的长度。

3）在空间上传播背景像素样本：相邻的背景像素共享相似的时间分布，并且像素的新背景样本也应该更新相邻像素的模型。 根据这一政策，由前景隐藏的背景模型将不时地从相邻像素位置更新背景样本。 这允许关于依赖于专门作为背景的样本的背景进化的信息的空间扩散。 因此，我们的背景模型能够适应不断变化的照明和结构演变（添加或删除的背景物体），同时依靠严格的保守更新方案。

更准确地说，让我们考虑像素x的4或8连接的空间邻域，即NG（x），并假设已经决定通过插入v（x）来更新样本集M（x）。 然后，我们还使用这个值v（x）从附近的一个像素根据均匀定律随机选择的样本集M（y∈NG（x））来更新。

由于像素模型包含许多样本，可能意外插入邻居模型的无关信息不会影响检测的准确性。 此外，不相关信息的错误扩散被阻止，因为在进一步传播之前需要匹配观察值。 这种自然限制抑制了误差的扩散。

## 3.2 改进及方法描述

3.2.1基于ViBe算法的改进

ViBe算法的具有很多优势：快速抑制，相机抖动的内在韧性，噪音弹性和嵌入式适用性。但是在背景初始化时采用像素样本集，容易产生Ghost区域。本文提出了三个方面的改进。

1. 不同的距离函数和阈值准则：摒弃原来的几何计算距离方法，使用圆锥模型。几何计算距离公式（6）：

（6）

公式8计算方便简单，但是灰度特征单调，RGB的颜色空间较多，空间模型如图10所示，像素点在RGB颜色空间的圆锥模型计算距离为公式（7）

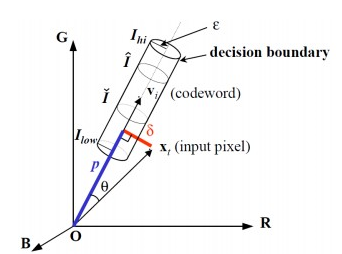


图10



 （7）



颜色距离可以通过下列公式计算：





Ri，Gi，Bi表示像素点在R，G，B三个通道的值，三个通道之间存在很大的相关性，使运动目标的检测和切割更加方便。

通过样本集的方差决定阈值，阈值与样本集的方差成比例，样本集的方差越大，背景越复杂，判断阈值越大。假设样本集方差为，设置一个阈值0.5，范围在20~40之间，检测效果良好。

2）使用不同大小的形态学方法

采用目标整体的观点，以补偿基于像素的前景检测的缺点。对于替代掩码和分割掩码使用不同大小的形态学方法来提高检测精度。实验结果如图12所示：



（a）原图 （b）基于ViBe分割效果



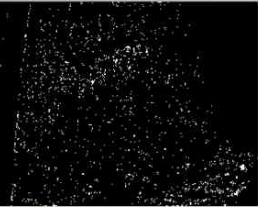
（c）采用改进方法分割效果

图12.对比效果图

由图12（a）可以看出，一些背景像素会随机进入到像素模型中，前景像素也成为了背景像素，该方法对抑制Ghost也非常有效，减少了背景像素出现在前景目标中。

3）检测闪烁像素点

引入了闪烁的概念。 当像素的更新标签与前一帧的更新标签不同时，闪烁级别增加15，否则将其递减1，然后根据大小来确定像素的闪烁点 闪烁的水平。 闪烁的像素主要出现在诸如叶子，水线等复杂的背景场景中。这些场景将出现频繁变化的像素背景和前景中，因此这些闪烁应该单独处理，可以作为一个整体用作背景。使用此种策略，可以减少抖动背景序列的失误检测。实验结果如图13所示：



（a）原图 （b）基于ViBe效果



（c）采用改进方法效果

图13.对比效果图

由图13可以看出，叶子的抖动让视频背景变得比较复杂，ViBe算法中有大量背景像素点不被希望的出现在了前景检测中，改进后的算法对这种抖动的现象抑制了大量的背景像素点。

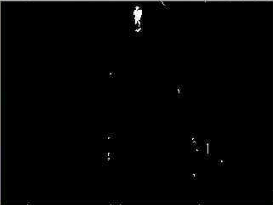
4）增加更新因子

ViBe算法中，默认更新因子为16，当背景变化很快时，背景模型不能快速更新，会导致前景检测中出现更多错误。 因此，有必要根据背景变化的程度来调整更新因子的大小，更新因子可以分为几个等级，如rate = 16，rate = 5，rate = 1。

由图14可以看出，使用默认因子16时会出现大量背景像素点，调整更新因子可以去除大量这样的像素点，同时不影响运动目标的形状。



（a ）原图 （b）基于ViBe效果



（c）采用改进方法效果

图14.对比效果图

3.2.2再浓缩方法

采用改进的ViBe算法，完成了视频浓缩，效果良好。在保证有效帧不丢失基础上进行进一步的浓缩。考虑客观的可定量的特征相识度来评判，在现有的浓缩视频的基础上按帧定长间隔抽取帧而形成更浓缩的视频，设置一个步长，抽取帧与步长内的帧进行对比，在一个步长内，视频场景变化很小，内部相似度很高，一头一尾就可以描述这个步长内视频的内容，去除相似度高的帧。设定一个相似度差异的阈值，在相似度差异的阈值范围内，找到更浓缩的视频序列。

步长和阈值可以自适应地调整，如果抽取帧与步长内的帧相似度都很高，就扩大步长，否则就缩小步长。

相似度比较这里采用了简单快速的感知哈希算法，基本思想是为每个图像生成“指纹”字符串，然后比较不同图像的指纹。 结果越近，图像越相似。为了保留亮度、框架等信息，忽略每一帧图片的细节，将图片缩小。简化缩小的图片色彩转为灰度图像并计算所有灰度像素值的平均值。再将每个像素灰度值与平均值进行对比。计算哈希值，将上一步的结果组合在一起就构成了一个数作为图片的“指纹”。通过计算“汉明距离”，不同的位数小于5则图片相似。算法如图12所示：



图12.算法流程图

实验结果表明，原浓缩视频429帧，播放视频17秒，经改进后浓缩视频313帧，播放视频13秒。

# 第四章 原型系统设计与实现

## 4.1 开发环境及工具简介

本次系统设计在VC++环境下，开发工具使用了Microsoft Visual 2013和应用了OpenCv视觉类库。使用了C++中的类、封装、继承、多态性等机制。OpenCv是Intel开源计算机视觉库，采用C/C++语言编写，实现了图像处理和计算机视觉方面的很多通用算法。系统视频处理过程中直接运用OPENCV基本数据类型，提高系统有效性。

## 4.2 项目实施的对环境和社会的影响

随着4G的广泛应用和视频的大规模生产，视频浓缩将成为社会发展的需要。实验表明，人们不能长时间专注于单调的事物，如果人眼注视视频屏幕超过22分钟，那么将会有95％的屏幕信息被忽略[j]。可以看出，绝大多数视频捕获只能起到后验证的作用。而查证通过人播放视频逐帧进行查看，没有情节的视频序列的情节很难让人专注，大量视频观看是一件非常困难的事情。保证不丢失移动目标信息中的情况下，进行有效的浓缩，快速捕捉视频内容和事件，可适用于法律，秩序，交通，刑事调查等警务业务。视频浓缩系统将大大提高应急响应能力的实施，使视频信息能够发挥最大的效率。

视频浓缩在提取视频帧和智能分析的基础上更符合用户的需要。 视频浓缩系统将是未来的研究趋势，视频浓缩的检索系统将成为安全监控和网络视频研究热点领域。在安全监控领域，将获得获得了广泛的应用，从而激励更多的创业公司投身于视频浓缩领域。

## 4.3 系统设计

4.3.1 系统总框架类的设计

借助OpenCV提供的视频图像处理库，实现改进的ViBe算法，逐帧对视频进行背景建模，通过背景差法，判断每一帧是否含有感兴趣的前景对象，如果不存在，则将该视频帧略去，只保留含有前景对象的视频帧，最后将保留下来的视频帧整合成新的视频文件，达到视频浓缩的效果。在现有浓缩视频的基础上，采用感知哈希算法来比较每一帧图像的相似度。形成更加浓缩的视频。系统总流程图如图13所示：

图13.系统总流程图

4.3.2 系统各模块设计

1. 读取视频文件

对不同格式视频的序列文件的读取是通过OpenCv提供的函数库实现的。 对于图像或视频的加载、显示、窗口设置、工具条、图像合并为视频、视频分割成图像、打开视频或摄像头等操作的定义是放在OpenCV的highGui模块中的。OpenCV中有一个VideoCapture类专门读取摄像头或者视频文件的类。

2）对视频序列进行背景建模。

在按帧读取视频文件后，通过ViBe算法完成视频背景模型的建立，并对之后每一帧进行更新。流程图如图16所示：

（1）背景模型初始化

ViBe 采用单帧初始化背景模型，根据相邻像素具有相似临时分布的假设，可采用像素空间邻域分布信息构建背景模型，即随机选取邻域值初始化背景模型。邻域窗口的选择需满足两个条件：包含充分的样本，距离计算使用RGB模型。这种初始化方式使运动目标的检测和切割更加方便。

（2）背景模型更新策略[8]

背景模型的更新是运动目标检测算法的关键，ViBe 采用一种指数单调衰减的方式保证样本的生命周期呈单调衰减，并非传统的先进先出替换策略，每个样本被选中的概率是相等的。改进ViBe 的更新策略包括三个部分：对于替代掩码和分割掩码使用不同大小的形态学方法来提高检测精度；加入检测闪烁像素点，可以减少抖动背景序列的失误检测；调整更新因子可以去除大量这样的像素点，同时不影响运动目标的形状。



图16.背景建模流程图

3）对视频帧进行划分。

通过背景差法[f]，准确的定义前景并精确的提取出前景，在前面建立背景模型的基础上，并用该模型不断的与视频中的每一帧图像进行比较，图像中与背景相似的区域被认为是背景，当前最新的图像与背景做差，即可求得背景差图，然后对该图进行二值化，最终获得运动前景区域，最后利用该图像信息更新背景模型，供下次检测使用。流程图如图17所示：



图17.提取前景对象流程图

4）重新整合视频文件。

利用OpenCV的视频图像库整合有效帧。首先生成视频流对象，然后将含有感兴趣对象的视频帧添加进视频流，将不含有感兴趣对象的视频帧略去，最终生成浓缩后的视频文件。流程图如图18所示：



图18.整合视频流程图

5）再浓缩视频

保证有效帧不丢失基础上进行进一步的浓缩。考虑客观的可定量的特征相识度来评判，在现有的浓缩视频的基础上按帧定长间隔抽取帧而形成更浓缩的视频，设置一个步长，抽取帧与步长内的帧进行对比，在一个步长内，视频场景变化很小，内部相似度很高，一头一尾就可以描述这个步长内视频的内容，去除相似度高的帧。设定一个相似度差异的阈值，在相似度差异的阈值范围内，找到更浓缩的视频序列。流程图如图19所示：



图19.再浓缩视频流程图

## 4.4 系统实现及测试

本小节主要介绍几个核心类的设计。

4.4.1 界面设计

CMFCDemoDlg类功能函数描述如下：

1. CMFCDemoDlg：构造函数；
2. OnInitDialog：生成消息映射函数，处理消息程序；
3. OnBnClickedOk：选择对话框；
4. OnQueryDragIcon：用户拖动最小化窗口时系统调用此函数取得光标显示；
5. OnPaint：是整个界面工作区居中；

（6）OnBnClickedButton1：打开视频，获取视频所在位置信息，点击后可以播放视频；

（7） OnBnClickedButton2：视频处理，点击后进行背景建模，显示前景信息。

|  |  |
| --- | --- |
| 类名 | CMFCDemoDlg |
| 数据成员 | HICON m\_hIcon;  enum { IDD = IDD\_MFCDEMO\_DIALOG }; |
| 成员函数 | CMFCDemoDlg(CWnd\* pParent = NULL);  virtual void DoDataExchange(CDataExchange\* pDX);  virtual BOOL OnInitDialog();  afx\_msg void OnSysCommand(UINT nID, LPARAM lParam);  afx\_msg void OnPaint();  afx\_msg HCURSOR OnQueryDragIcon();  DECLARE\_MESSAGE\_MAP()  afx\_msg void OnBnClickedOk();  afx\_msg void OnBnClickedButton1();  afx\_msg void OnEnChangeEdit2();  afx\_msg void OnBnClickedButton2(); |



[a] <http://www.authenmetric.com/msg.php?id=13>

[b]王海军，葛红娟，张圣燕 《基于L1范数和最小软阈值均方的目标跟踪算法》

[c]<http://briefcam.com/>

[d]Sebastian Brutzer, Benjamin H oferlin, Gunther Heidemann

Intelligent Systems Group, University at Stuttgart, Germany.Evaluation of Background Subtraction Techniques for Video Surveillance

[e]http://www.cnblogs.com/ronny/archive/2012/04/12/2444053.html

[f]Olivier Barnich and Marc Van Droogenbroeck, Member, IEEE.ViBe: A Universal Background Subtraction Algorithm for Video Sequences

[j] 黄丹平.视频分析及管理系统在公安刑侦中的应用[J].中国安防.