



**毕业设计外文资料翻译**

**（译文）**

**题目名称：**基于OpenCV的动态视频监控及动作识别系统

**学 院：** 计算机科学技术学院

**专业年级：** 软件工程2019级

**学生姓名：** 董永泽

**班级学号：** 1班43号(转)

**指导教师：** 刘太辉

**二○二三年 三 月 一 日**

**译文题目：**Python和Opencv中的运动检测系统

**原文题目：**A Motion Detection System in Python and Opencv 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)

**原文出处：**2021Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)

# Abstract

This paper suggests a motion detection software system that enables us to see the movement around an object or a visual area. Motion detection software is widely preferred as one of the most important security features in recent days. It is used to enhance existing security devices such as motion sensor lighting on indoor and outdoor security cameras. Security cameras are one of the most common devices that use motion detection technology . Now, for better protection, motion detector cameras are preferred. The paper suggested here looks at one of the python motion detection software and detects the movement of objects in front of the camera and determines the time they stay close to the camera.

**Key words：**Motion Detection , OpenCV , Python , Surveillance , Bokeh.

Motion detection detects changes in object position concerning the environment and vice versa. This motion detection software helps us to detect moving objects in front of the camera . It can perform the following tasks using this software, such as Get a screenshot while working at home, Monitor your child before screen time, Get illegal access to your yard, Find unwanted community/animal movement around your room/house/tunnel and many other things. In this research work, a motion detection software is proposed which allows us to see the movement in front of the camera. Generally, there are many ways to do motion detection, tracking, and analysis in OpenCV, and in that some are very simple and some are very complex. Earlier Bayesian features (opportunities) are adopted that works based on the background and foreground data . All of these methods are concerned with separating the background from the front (and also provide ways to distinguish between actual movement and just a change of shade with a slight change in brightness). Now, why is this so important? And why do we care which pixels are lower and front and which pixels are the back part? With motion detection, we often make the following assumption: Our video streaming layer is very stable and does not change to consecutive video frames. So, if we can model the background, we monitor it to make a big change. If there is a major change, we can get it - this change is often similar to the movement of our video . Now it is clear in the real world that this belief can easily fail. Because of the dignity, thinking, lighting conditions, and other possible environmental changes, our background can look very different from the various video frames. And if the background looks different, it can discard our algorithms. This is why effective retrieval / pre-discovery systems use horsemounted cameras and controlled lighting conditions. In the proposed software program, initially, it captures the first frame via webcam. This framework will be treated as the first framework. The movement will be obtained by calculating the phase difference between the first frame and the new frame. The new frames will be called Delta\_frame, in which you will improve your delta frame using pixel power. The refined structure will be called the Threshold framework. Then by using other complex image processing methods such as Shadow Removal, Dilation, Contouring, etc. in the threshold framework larger objects are captured . Also, it could be possible to capture the timestamp when the object is framed and out of frame. Therefore, it is possible to get the screen in time. Finally, a graph will be created that will show us how long an object stays in front of the camera and how often the moving object is detected.

In Roquiero and Petrushin developed a system that could calculate the natural population with the help of a camera. They do this by analyzing the images obtained and removing certain features with the help of motion detectors. This is done using a medium filter that took several images at a time and calculated the average value of each pixel. After that the middle image is subtracted in pixels by pixel from the recent image, if the difference is higher than a certain threshold is marked as before. Front pixels are areas that have moved overtime. Later they used a learning model of machines that are trained to count people in pictures on previous pixels. By following those steps they reached 98.64% accuracy in detecting if there were more than one person watching. They achieved an 86.9% accuracy in calculating the exact number of people in a photograph using the closest neighbor algorithm. Processing and pre-extracting techniques used in this paper are used in our application. In the authors try to create a good system for finding movement in sensitive areas such as banks and business ventures. They use the same method as stated in. First they take sample photos on webcam recording images, and then store those images in a bar that they use to calculate the difference between the sample images. Each time they saw the movement of the counter increasing and when it was moving above a certain threshold they sent a text message to the cell phone that the movement was happening and the bus was opened. This paper outlines some of the image processing methods we have chosen to use in our solution. In Wei, Li, Yue introduced a front-to-back algorithm used for motion detection for the purpose of tracking objects moving from a video camera. The authors have achieved this by supporting their solution in the Gaussian Mixture Model (GMM), which is a work in progress. Authors modified the traditional GMM algorithm by adjusting the Gausian parameters and number of parts, the authors did this to reduce computer time. Authors used their GMM algorithm for each pixel in each frame to perform pre-domain classification. Once the front and background have been defined the authors make the frame binary, while the front is white (moving objects) and the back is black (no moving objects). The authors then take binary pictures with filters to remove sounds such as wind, light and shadows. The authors compare their GMM algorithm with the traditional GMM algorithm using a test in which a video camera tracks moving objects on traffic and compass. The result was that traditional GMM had 58% accuracy while their GMM algorithm had 77% accuracy of tracking traffic. The solution used in this paper is considered to be the decision of how to obtain movement.

In Tahkker and Kapadia used OpenCV on Android. Their goal was to give students an idea of what it takes to run OpenCV on Android. The authors demonstrated the effects of applying certain filters to an image, such as color change, gray scale and dilation. The authors also tested the features on three different devices to check the time it took for the algorithms to complete. In the authors developed a tutorial guide on the concept behind how to use OpenCV for image recognition purposes. The article covers image management and a variety of machine learning capabilities that contain OpenCV. The first part of the paper explained how to handle images in the form of filters, object recognition and image classification. The second part explained how different machine learning algorithms work as decision trees, k-mean combinations and neural networks. This document helped you understand how to use and use OpenCV to create an application that can detect human availability.

To make this software we need libraries of OpenCV, Bokeh, Pandas, Datetime. OpenCV-Python is a library of Python bonds designed to solve computer vision problems. OpenCV-Python uses Numpy, which is a highly optimized library of pricing performance with the syntax of MATLAB style. All properties of the same OpenCV members are converted and removed from the same Numpy members. This also makes it easier to integrate with other Numpy libraries such as SciPy and Matplotlib. Bokeh is an interactive library for modern web browsers. It provides beautiful, concise architecture for flexible graphics, and provides highperformance connections to large databases or streams. Bokeh can help anyone who would like to quickly and easily create interactive sites, dashboards, and data applications. In a computer program, pandas is a software library written in Python programming language for cheating and analyzing data. In particular, it provides data structure and functionality to manipulate numerical tables and time series. It is a free software licensed under the three-term BSD license.

Initializing the variables and capturing the video frames : We will be using these variables in the code and better understand them ,first\_frame , satus\_list , times ,df for dataframe. OpenCV has in-built functions to open the camera and capture video frames. “0” denotes the camera at the hardware port number 0 in the computer. We are capturing the video frames in a new variable , video.

Converting the captured frame to gray-scale : We change the color frame to a gray frame as an additional layer of color is not required. GaussianBlur is used to smooth the image and will improve the accuracy of detection [13]. These are set with higher order calculation theorems, so you can use standard kernel values as (21,21) and standard deviations as 0.

Capturing The Frames : The first frame will be treated as the first frame. Motivation will be obtained by calculating the phase difference between this basic framework and the new frames that contain something. Therefore, we use the absdiff function and call a different frame that emerges as a delta framework [14]. As a measure, you can select 30 pixels as the limit value, and define the color of the limit value to be white, the Binary threshold function is a continuous function that works with 2 different values: either 0 or 1. We view the current frame status as 0 when nothing is present in front of the camera or as 1 when an object is in front of the camera.

Contouring and Threshold Dilation : Each part of the object casts shadows on the back, or on other parts of it. This may seem confusing.To reduce these types of sounds, we need to filter the image. In the Dilate function, we can set the smoothness level by setting the multiplication value. The more the repetitive value increases, the smoother it will be, the more time it will be to process. Once the frame is filtered, we will have to get the look of our frame. A line turns when a function has a constant value in all points. We need a view of our current frame to identify the size and location of the object. To achieve this, we transfer a filtered frame copy to the FindCountors process. We use a copy of the filtered frame to get the drawings not the original ones as we do not want to disturb the original filtered frame. This can be customized using the concept of a peer location. Here, we skip those objects with an area of less than 10,000 pixels. In the larger area lines, we set the state = 1, that is, the object is available.

A simple and efficient motion detection system is proposed in this research work. This project supports the user to interact with the machine and made possible to access and extract information from the internet and our PC. The proposed model is useful due to its effective and user-friendly nature. It is technically free to use and very helpful in today’s technologydependent generation ensuring their security. Also, in a world that is full of technological advancements that ensure human comfort, a webcam motion detector is adding itself to this race. It could be extended further and improvised as per the user’s requirements

# Python和Opencv中的运动检测系统

摘要

本文提出了一种运动检测软件系统，使我们能够看到物体或视觉区域周围的运动。运动检测软件作为最近几天最重要的安全功能之一被广泛首选。它用于增强现有的安全设备，如室内和室外安全摄像头上的运动传感器照明。安全摄像头是使用运动检测技术的最常见设备之一。现在，为了更好地保护，运动检测器相机是首选。本文建议查看一款python运动检测软件，检测相机前物体的运动，并确定它们靠近相机的时间。

**关键字：**运动检测；OpenCV；Python；Surveillance; Bokeh

运动检测检测与环境有关的物体位置的变化，反之亦然。这个运动检测软件可以帮助我们检测摄像机前的运动物体。它可以使用该软件执行以下任务，例如在家工作时获取屏幕截图、在屏幕时间前监控您的孩子、非法进入您的院子、在您的房间/房子/隧道周围发现不需要的社区/动物活动以及许多其他事情。在这项研究工作中，提出了一种运动检测软件，它可以让我们看到相机前面的运动。通常，在OpenCV中有很多方法可以进行运动检测、跟踪和分析，其中有些非常简单，有些非常复杂。早期的贝叶斯特征（机会）被采用，其基于背景和前景数据工作。所有这些方法都与将背景与正面分离有关（并且还提供了区分实际移动和只是亮度略有变化的阴影变化的方法）。为什么这这么重要？为什么我们关心哪些像素在下面和前面，哪些像素在后面？对于运动检测，我们经常做出以下假设：我们的视频流层非常稳定，不会更改为连续的视频帧。因此，如果我们能够对背景进行建模，我们就会对其进行监控，以做出重大改变。如果有重大变化，我们可以得到——这种变化通常与我们视频的运动相似。现在很明显，在现实世界中，这种信念很容易失败。由于尊严、思维、照明条件和其他可能的环境变化，我们的背景看起来与各种视频帧非常不同。如果背景看起来不同，它可以丢弃我们的算法。这就是为什么有效的检索/预发现系统使用骑马的摄像机和受控的照明条件。在提议的软件程序中，最初，它通过网络摄像头捕捉第一帧。这一框架将被视为第一个框架。移动将通过计算第一帧和新帧之间的相位差来获得。新的帧将被称为Delta\_frame，其中您将使用像素功率改进您的Delta帧。细化后的结构将被称为阈值框架。然后通过在阈值框架中使用其他复杂的图像处理方法，如阴影去除、扩张、轮廓等，捕获更大的对象。此外，当对象被框起来和超出框时，可以捕获时间戳。因此，可以及时获得屏幕。最后，将创建一个图形，向我们显示物体在相机前停留的时间以及移动物体被检测到的频率。

在Roquiero和Petrushin开发了一个系统，可以在相机的帮助下计算自然种群。他们通过分析获得的图像并在运动检测器的帮助下去除某些特征来做到这一点。这是使用一个介质滤波器来完成的，该滤波器一次拍摄多个图像并计算每个像素的平均值。之后，如果差值高于某个阈值，则从最近的图像中逐个像素地减去中间图像，标记为如前所述。前面的像素是随着时间推移而移动的区域。后来，他们使用了一种机器的学习模型，这些机器经过训练，可以在以前的像素上对图片中的人进行计数。通过遵循这些步骤，他们在检测是否有不止一个人在观看时达到了98.64%的准确率。他们使用最近邻算法计算照片中的确切人数，准确率达到86.9%。在我们的应用中使用了本文中使用的处理和预提取技术。在这篇文章中，作者试图创建一个良好的系统，用于在银行和企业等敏感领域寻找行动。他们使用与中所述相同的方法。首先，他们在网络摄像头上拍摄样本照片，记录图像，然后将这些图像存储在一个栏中，用于计算样本图像之间的差异。每次他们看到计数器的移动增加，当它移动到某个阈值以上时，他们都会向手机发送一条短信，说明移动正在发生，公交车被打开了。本文概述了我们选择在解决方案中使用的一些图像处理方法。在魏，李，岳介绍了一种用于运动检测的前到后算法，用于跟踪摄像机中移动的物体。作者通过在高斯混合模型（GMM）中支持他们的解决方案实现了这一点，这是一项正在进行的工作。作者通过调整高斯参数和零件数量来修改传统的GMM算法，以减少计算机时间。作者对每帧中的每个像素使用他们的GMM算法来执行域前分类。一旦定义了正面和背景，作者就将帧设为二进制，而正面是白色（移动对象），背面是黑色（没有移动对象）。然后，作者用滤镜拍摄二进制图片，以去除风、光和阴影等声音。作者使用摄像机跟踪交通和指南针上的运动物体的测试，将他们的GMM算法与传统GMM算法进行了比较。结果是，传统的GMM算法对流量的跟踪准确率为58%，而他们的GMM方法对流量的追踪准确率为77%。本文中使用的解决方案被认为是如何获得运动的决定。

在Tahker和Kapadia在Android上使用了OpenCV。他们的目标是让学生了解在Android上运行OpenCV需要什么。作者演示了对图像应用某些滤镜的效果，如颜色变化、灰度和膨胀。作者还在三种不同的设备上测试了这些功能，以检查算法完成所需的时间。在中，作者开发了一个关于如何使用OpenCV进行图像识别的概念的教程指南。本文介绍了图像管理和包含OpenCV的各种机器学习功能。论文的第一部分介绍了如何以滤波器、物体识别和图像分类的形式处理图像。第二部分解释了不同的机器学习算法如何作为决策树、k均值组合和神经网络工作。本文档帮助您了解如何使用OpenCV创建一个可以检测人员可用性的应用程序。

为了制作这个软件，我们需要OpenCV、Bokeh、Pandas、Datetime的库。OpenCV Python是一个Python债券库，旨在解决计算机视觉问题。OpenCV Python使用Numpy，这是一个高度优化的定价性能库，具有MATLAB风格的语法。相同OpenCV成员的所有特性都将转换并从相同Numpy成员中删除。这也使它更容易与其他Numpy库（如SciPy和Matplotlib）集成。Bokeh是一个用于现代网络浏览器的交互式库。它为灵活的图形提供了美观、简洁的体系结构，并提供了与大型数据库或流的高性能连接。Bokeh可以帮助任何想要快速轻松地创建交互式网站、仪表板和数据应用程序的人。在计算机程序中，pandas是用Python编程语言编写的用于欺骗和分析数据的软件库。特别是，它提供了操作数值表和时间序列的数据结构和功能。它是一个根据BSD三期许可证授权的自由软件。

初始化变量并捕获视频帧：我们将在代码中使用这些变量，并更好地理解它们，数据帧的first\_frame、satus\_list、times、df。OpenCV具有打开相机和捕捉视频帧的内置功能。“0”表示计算机中硬件端口号为0的摄像头。我们正在一个新的变量video中捕获视频帧。

将捕获的帧转换为灰度：我们将彩色帧更改为灰色帧，因为不需要额外的颜色层。GaussianBlur用于平滑图像，并将提高检测的准确性[13]。这些都是用高阶计算定理设置的，所以你可以使用标准核值作为（21，21），标准偏差作为0。

捕捉帧：第一帧将被视为第一帧。通过计算这个基本框架和包含某些内容的新框架之间的相位差，可以获得动机。因此，我们使用absdiff函数并调用一个不同的框架，该框架作为delta框架出现[14]。作为一种度量，您可以选择30个像素作为极限值，并将极限值的颜色定义为白色，二进制阈值函数是一个连续函数，可使用2个不同的值：0或1。当相机前面什么都不存在时，我们将当前帧状态视为0，当对象在相机前面时，我们视为1。

轮廓和阈值扩展：物体的每个部分都会在背面或其他部分投射阴影。这可能看起来很令人困惑。为了减少这些类型的声音，我们需要过滤图像。在Dilate函数中，我们可以通过设置乘积值来设置平滑度。重复值增加得越多，就越平滑，处理的时间就越多。一旦帧被过滤，我们就必须获得帧的外观。当一个函数在所有点上都有一个常数值时，线就会转动。我们需要当前帧的视图来识别对象的大小和位置。为了实现这一点，我们将过滤后的帧副本传输到FindCountors进程。我们使用过滤框架的副本来获得图形，而不是原始图形，因为我们不想干扰原始过滤框架。这可以使用对等位置的概念来定制。在这里，我们跳过那些面积小于10000像素的对象。在较大的区域线中，我们将状态设置为1，即对象可用。

本文提出了一种简单高效的运动检测系统。该项目支持用户与机器交互，并使其能够从互联网和我们的电脑中访问和提取信息。由于其有效且用户友好的特性，所提出的模型非常有用。它在技术上是免费使用的，在当今依赖技术的一代人中非常有帮助，确保了他们的安全。此外，在一个充满技术进步以确保人类舒适的世界里，网络摄像头运动检测器正在加入这场比赛。它可以根据用户的要求进一步扩展和即兴制作。