

National Research University Higher School of Economics

Faculty of Business Informatics

School of Software Engineering

Software Management Department

AN APPLICATION FOR DYNAMIC OBJECT IDENTIFICATION BASED
ON LUCAS-KANADE ALGORITHM

Student: Kostenko Dmitry

Group: 472SE

Argument Consultant: Prof. Ivan. M. Gostev, PhD

Style and Language Consultant: Tatiana A. Stepantsova

Moscow

2013

Abstract

In this paper we present an approach of object identification and tracking. We use a differential method for optical flow calculation developed by B. D. Lucas and T. Kanade [1]. The proposed approach consists of three steps which can be executed in real-time. As a first step the 2D vector of optical flow is calculated by Lucas-Kanade method [1]. Then it produces a binary vector of the optical flow vector to generate regions that are moving. Finally, program divides moving regions of objects, using differences in speed. To demonstrate the moving objects, they can be marked with different colors.

Содержание

Introduction	4
Problem statement	6
Algorithm	8
Related work	10
Conclusion	11

Introduction

In our generation scientific and engineering community attempts to build human-like machines. Like a man, these robots are supposed to have various senses, ability to analyse incoming information, based on its conclusions and develop and implement behaviour patterns. And starting to implement, scientists immediately came across a problem of image understanding. There are a lot of engineering achievements that supersede analogues of the most perfect system - human. However, there is a large gap between human and machine in the technology of artificial intelligence. Impossibility to fully automate analysis of information from visual channel, even as a child level, pushes researchers to move forward gradually. They divide the problem into sub-problems of computer vision. For example, improving image quality, identifying features points and determining the similarity of images.

The vision channel is one of the most informative ones. Data volume from video stream exceeds by several times volumes from other sensors. Here lies the pitfall - redundancy of information. Occasionally just a few bytes of information are enough. For example, we need only the object of interest, and in addition to this, we have other objects, background and fine detail (an object of interest is what, which we need to track). In contrast to human, the machine can not optimally handle this task. Consequence of this problem is the fact that even an ant can navigate better in the complicated situation than existing robots.

Why is the task of understanding the image so complex? In the image processing we can come across several problems. Firstly, objects of observations are very volatile in general. Causes of variability are difficult to formalize. Existing methods of dealing with different lighting, noises and distortions cope rather clumsily. Secondly, scene observations also can not be modelled. This is due to the fact that there is a diversity of geometrical forms, colors and textures. Third problem is knowledge base. Any system required to recognize an object should be aware of it. It means that system should be trained to recognize images. People, during their life, spend a lot of time doing it.

Nowadays, paradigm proposed by D. Marr [2] is used in image processing. Its point is to process images in sequence:

1. image pre-processing
2. segmentation
3. distinguishing of the geometric structure
4. determining the semantics

In other words, processing of low-level should be performed at first, then intermediate-level processing and finally high-level processing.

In the general form, motion analysis is a comparison of consecutive images of the observed scene, which is the difference between the current frame and the previous frame. As the result of this operation, we have contours of a moving object. This is one of the basic ideas of our work.

The rest of this paper is organised as follows:

Section 2

Identification of key problems of object tracking and problem statement.

Section 3

An algorithm for solving the problems and its implementation.

Section 4

Related work.

Section 5

Summing-up, prospects of our algorithm and formulating the scope of our method.

Problem statement

One of the most important task in video observation consists of detection of objects of interest and tracing its mechanical trajectory in the next frames. This task can be broken down into sub-tasks. Firstly, point out the object of interest. After decision, we have the coordinate position of the object. On how to point out the object of interest, rectangular frame or pixel mask, will be written further. And then, we need to keep track of it on all subsequent frames.

Movement of objects in front of camera or camera movement in the fixed surroundings leads to changes in the image. The image of visible object movement is called optical flow [3]. There are a lot of methods for calculating optical flow. Lucas and Kanade [1] proposed differential approach. It is based on finding of velocity of each image's pixel. For calculating optical flow several suggestions were made.

1. the image is a continuous function of two variables
2. brightness of the object remains constant in a short period of time
3. tracing object on new frame in a short distance from itself on old frame

The first suggestion gives us all existing methods of mathematical analysis and allows to perform all mathematical operations on images. It is expressed by the equation: $\frac{\delta(x,y,t)}{\delta t} = 0$ The second one is because we live in the real world, so object can not travel great distances instantaneously. And the third suggestion exist only because without it we will not be able to find tracing object.

Based on these assumptions, first methods were based on calculating first-order derivative.

Одна из основных задач видеонаблюдения - состоит в выделении в видео объектов интереса и отслеживание траектории их движения в последующих кадрах. 2 задачи
1) выделение объектов интереса как выделить? прямоугольная рамка попиксельная маска

2) после решения первой задачи есть координаты положения объекта нам нужно отследить его на всех последующих кадрах

что такое объекты интереса - это все, что мы хотим отслеживать

проблемы на пути создания системы видеонаблюдения: главная проблема - это то, что видео очень большое по размеру. на видео может быть больше, чем один объект, поэтому нагрузка может быть очень высокая

вторая проблема - вид объекта от кадра к кадру будет меняться из-за изменения освещения, изменения позы, или иных изменений. (человек и развивающийся плащ)

третья проблема (про мышей) на видео могут быть несколько похожих друг на друга объектов, которые могут еще и перекрывать друг друга. задача отслеживания нескольких мышей в одной клетке до сих пор полностью не решена. т.к. они могут налезать друг на друга и после того, как они расползутся распознать какая из них какая абсолютно невозможно.

Самое простое решение задачи выделения фона - это вычесть из изображения с объектом изображение без объекта в результате мы получим изображение на котором будет явно выражен объект

Algorithm

Для пояснения алгоритма возьмем простую ситуацию.

Мышь ходит в банке.

Камера снимает сверху, что происходит в банке.

Будем считать, что изображение состоит из двух частей: фон и передний план.

До того, как опустим мышь в банку сфотографируем пустую банку.

В результате мы получим чистый фон.

Так как мы будем использовать стационарную камеру, то для получения переднего плана надо вычесть фон из каждого последующего изображения.

Эта разница и будет предположительно мышью.

Поскольку у камеры есть какой-то шум, то в результате вычитания мы можем получить лишние пиксели, которые не могут являться мышью.

Есть 2 метода борьбы с этим явлением.

1) Обработать изображение фильтром шума (например фильтром Гаусса) 2) Ввести некоторый порог.

Если разница больше либо равна порогу, то пиксель принадлежит переднему плану.

Если разница меньше порога, то пиксель принадлежит фону.

Применяя оба случая в нашей системе, мы все-таки можем получить шум на изображении.

Поэтому после вычитания применим методы математической морфологии для устранения шумов.

Например, сначала эрозию, потом дополнение.

Далее на таком бинарном изображении мы можем запустить выделение связанных компонент.

В результате каждая связанная компонента будет являться искомым объектом.

В нашем случае мышью.

Естественно, что такой алгоритм дает удовлетворительный результат только в искусственных условиях, т.е. когда у нас не меняется фон в результате изменения освещения.

В реальных условиях фон существенно меняется с течением времени.

Например камера видеонаблюдения, установленная на улице, днем и ночью будет иметь существенно различные изображения, а соответственно мы это должны учитывать в построении нашей системы видеонаблюдения.

Существует несколько способов получения изображения фона в реальных условиях:

1) Метод усреднения фона.

Усреднить попиксельное несколько последовательных кадров.

Данный метод неудовлетворительно работает, когда в кадрах есть движущиеся объекты или резкие изменения яркости.

2) Метод медленного фона.

В результате применения данного метода мы будем иметь более качественную модель фона.

Но у него есть свои дефекты при вычислении фона медленным методом.

Например, во время захвата фона в кадре находились 2 объекта в неподвижном состоянии, то они считаются частью фона.

Потенциально это может быть человек, который остановился, чтобы поговорить по телефону.

Если объект уйдет из наблюдаемой картины, то место, где он находился, будет считаться частью переднего плана.

После того, как мы выделили потенциальные объекты, тогда задача слежения за объектом сводится к задаче сопоставления объекта в текущем кадре с объектом в следующем кадре.

Возможно 3 случая состояния объектов на кадре:

- 1) новый объект
- 2) новое положение существующего объекта
- 3) исчезновение существующего объекта

Будем использовать детерминированный метод (метод полного перебора) для поиска нового положения объектов.

Он подразумевает под собой сопоставление ближайших объектов, используя предположения, описанные в п.3 (стр. ??).

Однако у него так же имеется ряд недостатков.

Например, применяя данный метод, в случае затора на автодороги, невозможно будет отличить одну машину от другой.

Но мы в нашей работе будем пренебрегать данным фактом и будем использовать метод сопоставления.

В результате имеем алгоритм видеонаблюдения:

- 1) обучение модели фона
- 2) для каждого нового кадра: - вычитание фона - обработка маски (фильтрация, математическая морфология) - выделение связных компонент - ассоциация объектов в треки - инициализация новых треков - обновление фона

Related work

1 page

Conclusion

1 page

Bibliography

- [1] Takeo Kanade Bruce D. Lucas. *An Iterative Image Registration Technique with an Application to Stereo Vision*. Computer Science Department Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, 1981.
- [2] D. Marr. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman and Company., 1982.
- [3] S. S. Beauchemin. The computation of optical flow. *ACM Computing Surveys*, 1995.