

Technische Universität Ilmenau
Fraunhofer Institut für Digitale Medientechnologie



Keyframe-Selektion zur Erhöhung der Erkennungsrate von Primaten in Videos

Masterarbeit zur Erlangung des akademischen Grades Master of Science

Wieland Morgenstern

Betreuer: Dipl.-Ing. Alexander Loos

Dipl.-Ing. (FH) Ronny Paduschek

Verantwortlicher Hochschullehrer:

Prof. Dr.-Ing. Dr. rer. nat. h.c. mult. Karlheinz Brandenburg,

FG Elektronische Medientechnik

Die Masterarbeit wurde am 12.05.2014 bei der Fakultät für Informatik und Automatisierung der Technischen Universität Ilmenau eingereicht.

Zusammenfassung

Die Ergründung des Verhaltens von Tieren ist für viele Forschungsbereiche wie Arten-erhaltung und Populationsüberwachung von hoher Bedeutung. Für eine umfassende Überwachung der Lebewesen werden häufig Videokameras eingesetzt. Die entstehenden Datenmengen können zur Entlastung des Menschen durch einen Computer analysiert werden und automatisch Informationen wie Spezies, Alter, Geschlecht und die Identität der einzelnen Tiere erkennen.

Diese Arbeit setzt es sich zum Ziel, die Erkennungsrate eines bereits vorhandenen Systems zur Identifikation von Primatengesichtern in Videos zu erhöhen. Es wird ein System implementiert, das aus einer Reihe von Aufnahmen desselben Individuums (einem Facetrack, der aus mehreren bis vielen hundert Frames bestehen kann) die besten Aufnahmen (Keyframes) extrahiert, um diese zur Identifikation zu verwenden. Dabei werden verschiedene Module implementiert, die Kriterien für die Auswahl aus den Frames ermitteln: Beleuchtung, Kontrast, Schwärzeanteil, Unschärfe, sowie die verschiedenen Varianten zur Erkennung der Pose des abgebildeten Tieres. Als beste Methode zur Posenerkennung stellt sich eine Klassifikation mit einer Support Vector Machine auf Gabor-Features heraus.

Aus den verschiedenen Parametern wird ein globaler Score berechnet, der zur Auswahl der Keyframes verwendet wird. Die Ergebnisse der Identifikation auf mehreren Keyframes können gewichtet miteinander verrechnet werden, um ein Gesamtergebnis zu erhalten. Gegenüber der bisherigen Variante, die Identifikation auf dem ersten Frame des Facetracks auszuführen, kann die Erkennungsrate deutlich gesteigert und unter bestimmten Umständen sogar verdoppelt werden.

Abstract

Wildlife monitoring plays an increasing role in many areas of research. Remote cameras and recording devices are used to estimate the size of populations and to monitor their behaviour. Computers can assist in interpreting and analysing the data that is collected. Information about species, age, sex as well as the identity of animals can be extracted automatically.

In this thesis we are trying to improve the recognition rate of an existing system that identifies great apes in video data. A system is implemented, which chooses the best (key) frames from a face track, a long (sometimes hundreds of frames) list of frames depicting the same individual. These key frames are then used in the identification process instead of the first frame, to boost the recognition rate.

Several modules are implemented to estimate different parameters in a frame: exposure, contrast, blackness, blur as well as an estimation of the head pose. The best variant of head pose estimation is a classification system that uses a Support Vector Machine on Gabor features of discrete poses.

A global score that describes the image quality is computed from the different parameters, which is then used to select the key frames. The identification results from several key frames can be put to a maximum voting to improve confidence in the results.

Compared to the previous approach of selecting the first frame for identification the recognition rate is greatly improved and under some circumstances, doubled.
