

Projektseminar I4: 'Angewandte Informationswissenschaft'

Andreas Schwanitz - 1945397

7. August 2016

Inhaltsverzeichnis

1	Allgemein	1
1.1	Welcher Themenbereich wird bearbeitet?	1
1.2	Welche Programmiersprache/Tools werden verwendet?	1
1.3	Wie setze ich mich im Projekt mit dem Thema auseinander?	2
1.4	Ist das Projekt zeitlich machbar?	2
1.5	Wo sind Eingrenzungen möglich?	2
2	Idee	2
2.1	Vorstellung	2
2.1.1	Beispiel	2
2.2	Konzept	2
2.3	Wie genau setze ich das Konzept um?	3
2.4	Worin liegt der Nutzen der Idee?	3

Zusammenfassung

1 Allgemein

1.1 Welcher Themenbereich wird bearbeitet?

Da ich mich darauf konzentriere eine neue Möglichkeit zur Datenanalyse zu schaffen, würde ich mich dem Bereich der Informatik zuordnen.

1.2 Welche Programmiersprache/Tools werden verwendet?

Als Programmiersprache werde ich Java in der aktuellen Version 8 verwenden. Welche Tools im Verlauf benötigt werden kann ich noch nicht ganz absehen.

1.3 Wie setze ich mich im Projekt mit dem Thema auseinander?

Ich werde mich eingehend mit zusammengeschnittenen Videos zu einem bestimmten Themenbereich auseinandersetzen und eine Software mit Datenbank entwickeln, die einzelne Ausschnitte aus einem solchen Video erkennen und zuordnen kann. Dies ermöglicht dann eine weitere Informatrische Analyse.

1.4 Ist das Projekt zeitlich machbar?

Zeitlich gesehen ist das Projekt, aus derzeitiger Sicht zeitlich machbar. Im Verlauf auftretende Probleme können natürlich nicht abgeschätzt werden. Da eine eigene Software für den Vergleich wie auch eine Datenbank mit Referenzinformationen erstellt werden muss ist die Anzahl der zu analysierenden Videos natürlich begrenzt.

1.5 Wo sind Eingrenzungen möglich?

Den Umfang kann man mit dem Themengebiet der zu bearbeitenden Videos eingrenzen wie auch mit der Anzahl der Videos bzw. mit der Datenbankgröße.

2 Idee

2.1 Vorstellung

Die Grundidee besteht darin, dass man eine Möglichkeit schafft um Videomaterial automatisch zu erkennen, auch wenn es aus unterschiedlichen Quellen kommt. Damit ist die Erkennung von verschiedenen Videos, die sich jeweils in Länge, Qualität und weiteren Merkmalen unterscheiden, aber auf ein Grundvideo referenzieren gemeint.

2.1.1 Beispiel

Als Beispiel seien Videos gemeint, die alle eine Szene aus dem Film 'König der Löwen' zeigen. Jedes Video hält eine Referenz auf den Film in sich, ist aber auf Grund von verschiedenen Längen, Qualitäten, Schnitten und Videoformaten nicht vergleichbar bzw. erkennbar.

2.2 Konzept

Mein Konzept basiert darauf, dass ich Perceptual-Hashing einsetze. Ich benutze diesen robusten Hashalgorithmus um Bilder aus den Videos zu extrahieren, normieren und diese dann schlussendlich in der Datenbank ab-

zulegen. Die Referenzinformation in der Datenbank kommen vom Originalvideomaterial. Für einen Vergleich wird nun der selbe Algorithmus auf das zu untersuchende Video angewendet und mit den Datenbankinformationen mittels Hammingdistanz verglichen.

2.3 Wie genau setze ich das Konzept um?

Ich schaue mich zunächst nach Videomaterial um. Ich schreibe eine Software, die Videomaterial einlesen und verarbeiten kann, sodass ich die Daten für die Datenbank erhalte. Ich Verleiche verschiedene Videos und evauliere die Ergebnisse.

2.4 Worin liegt der Nutzen der Idee?

Der Nutzen liegt klar auf der Hand. Bisher konnte man Videomaterial einzig an Hand seiner Metadaten wie dem Dateinamen oder ähnlichem erkennen. Mit dieser neuen Methode kann man den eigentlichen Inhalt eines Videos zur Analyse verfügbar machen und hat dadurch eine zweite Basis zur Kontrolle der bisher bekannten Daten und eine zusätzliche Quelle für weiterführende Analysen, die bisher nur Menschen zur Verfügung stand.