

Generarea metadatelor aferente conținutului video utilizând descriptori MPEG-7

- ABSTRACT -

Lucrarea abordează o tematica indexării și identificării datelor multimedia bazate pe conținut, precum și reprezentarea conținutului cu ajutorul metadatelor. Teza este alcătuită pe 6 capitole, în ultimele subcapitole s-a pus accentul pe contribuțiile personale și perspectivele ulterioare de dezvoltare a acestei tematici de cercetare.

Capitolul 1 realizează o introducere legată de tehnologiile de înregistrare digitală și apariția de noi posibilități de achiziție, stocare, schimb și distribuție a informațiilor audio-vizuale. Pentru facilitarea accesului utilizatorului la conținutul multimedia fără restricții tehnice, MPEG a dezvoltat un standardul MPEG-7 care permite descrierea eficientă d. p. d. v. semantic a conținutului multimedia și care facilitează interoperabilitatea.

Capitolul 2 al acestei tezei tratează noțiunea de metadata folosite pentru o descriere cât mai exactă și relevantă a conținutului multimedia utilizat. În cadrul descriptorilor video MPEG-7 sunt detaliați descriptorii de culoare, de textură, de formă, de mișcare, de localizare, precum și descriptorul de recunoaștere a fețelor.

Capitolul 3 plasează accentul pe metodele de segmentare video bazată pe regiuni: „region growing” și „split and merge”. În cazul simulării metricilor Minkovsky, Canberra și Cilindică, d. p. d. v. al diferenței maxime dintre imaginea originală și imaginea segmentată este evident că metrica Minkowsky are valorile cele mai mici. Sunt propuși doi algoritmi de segmentare spațială: algoritmul de segmentare cu prag al nuanței și cu prag variabil al nuanței. S-a obținut un număr mai mic de regiuni și folosind valorile pixelilor de start s-au determinat culorile dominante, procentul de apariție și ceilalți parametri necesari descriptorului de culori dominante (DCD).

În Capitolul 4 sunt prezentați algoritmi de estimare a mișcării obiectelor în secvențele video: FS, TSS, NTSS, SES, FSS, DS, OS, OTS, TDLS, CS, SS, ARPS. Algoritmii rapizi prezentați se apropie destul de mult, din punct de vedere al preciziei de estimare, de algoritmul de căutare completă. Diferența semnificativă apare în ceea ce privește numărul de blocuri verificate, iar din simulări rezultă RSZ de vârf este mai mare pentru algoritmi care verifică un număr mai mare de blocuri.

În Capitolul 5 au fost evaluate în Matlab performanțele în două scenarii diferite, realizându-se indexarea și identificarea imaginilor pe baza valorilor descriptorilor DCD. Din indexarea și identificarea celor 90 de imagini ale primului set s-a determinat o precizie medie de 30% pentru primele două distanțe de nesimilaritate și de 40% pentru a treia. Precizia a fost calculată în raport cu clasificarea imaginilor în grupuri în funcție de conținut. Din indexarea și identificarea celor 1000 de imagini ale setului 2 s-a determinat o precizie medie de 35% pentru primele două distanțe de nesimilaritate și de 73% pentru a treia. În acest caz deoarece imaginile nu erau clasificate în funcție de conținut s-a determinat precizia prin vizualizarea imaginilor rezultate și identificarea numărului de imagini relevante din primele 20 de rezultate.

Video Content Metadata Extraction using MPEG-7 Descriptors

- ABSTRACT -

This thesis approaches the theme of indexing and identification of multimedia content-based data together with content representation using metadata. The paper is organized in six chapters and the last paragraphs refer to personal contributions and further perspectives for development of this research theme.

Chapter 1 contains an introduction in digital record technologies and the possibility of appearance of new methods for acquisition, storage, exchange and distributions of audio-video informations. To allow user's access to multimedia content without any technical restrictions, MPEG created MPEG-7 standard. It permits an efficient description from semantic point of view of multimedia content and allows interoperability.

In the 2nd Chapter metadata are defined for a more precise and relevant description of multimedia content. From MPEG-7 descriptors there are detailed: colour descriptor, texture and shape descriptors, motion, location and face recognition descriptors.

Chapter 3 places the accent on region-based video segmentation: „region growing” and „split and merge”. When simulations for Minkovsky, Canberra and Cilindric metrics are performed it can be noticed from the point of view of maximum difference between original and spatial segmented image that Minkowsky metric leads to the smallest values. There are proposed two segmentation algorithms: hue threshold and variable hue threshold. Small number of regions there are obtained and using values of start pixels there have been determined dominant colours, appearance percentage and other parameters needed for DCD.

In Chapter 4 there are illustrated object-motion estimation algorithms in video sequences: FS, TSS, NTSS, SES, FSS, DS, OS, OTS, TDLS, CS, SS, ARPS. These fast algorithms approach, from estimation precision point of view, the complete search algorithm. The main difference is related to the number of selected blocks and from simulation the highest PSNR is achieved for those algorithms that select a large number of blocks.

Chapter 5 contains Matlab simulations for two different scenarios. Image indexing and identification based on DCD descriptors values are achieved. From indexing and identification of 90 images from the first set there is achieved a 30% medium precision for the first two dissimilarity distances and 40% medium precision for the 3rd one. The precision was determined based in images classification in groups depending on content. From indexing and identification of 1000 images of the 2nd set there is obtained a 35% medium precision for the first two dissimilarity distances and 73% medium precision for the 3rd one. In the second case the precision was determined by visualization of resulted images and identification of the most relevant ones out of 20 results (the images were not classified based on content).