

ХИЕРАРХИСКА КЛАСИФИКАЦИЈА НА ВИДЕО ПОДАТОЦИ

Ѓорѓи Мадаров, Ивица Димитровски, Сузана Лошковска, Дејан Ѓорѓевиќ

Универзитет „Св. Кирил и Методиј“ – Скопје, Факултет за електротехника и информациски технологии, Карпош II бб, П. Факс 574, 1000 Скопје, Македонија,
{madzarovg, ivicad, suze, dejan}@feit.ukim.edu.mk

Анстракт – Во овој труд е претставена хиерархиска архитектура од SVM класификатори за класификација на видео документи, базирана на претходно дефинирана класна хиерархија. Таа решава конкретен четирикласен проблем на класификација на видео податоци чија хиерархиска претстава се состои од две нивоа. Секој видео податок е опишан со обележја екстрахирани од автоматски детектираните клучни слики добиени со употреба на дескриптори од MPEG-7 стандардот. Архитектурата од SVM класификатори е базирана на бинарно класификациско дрво и е дизајнирана со цел да се постигнат подобри перформанси во однос на прецизност и брзината на класификација. Резултатите од експериментите покажаа дека оваа архитектура се одликува со голема прецизност особено во првото ниво на хиерархијата со грешка од само 6.6%.

Клучни зборови – Машини со носечки вектори (SVM), MPEG-7 дескриптори, хиерархиска класификација.

1. ВОВЕД

Содржински базираната класификација на видео податоците го решава проблемот на категоризација на видео секвенци во конкретни предефинирани видео жанрови. Поради огромниот број на лесно достапни дигитални видео библиотеки, пожелно е да имаме автоматска класификација и категоризација на видео податоците за крајниот корисник да може да изврши брзо и ефикасно содржинско пребарување. Проблематиката за содржински базирана класификација на видео податоци е интересно подрачје за истражување и е третирана од страна на неколку автори. Во своите истражувања тие нудат неколку различни пристапи. Ли (Li) користи Gaussian Mixture Model (GMM), како модел за класификација на видео податоците [1]. Дополнително користи и Principal Component Analysis (PCA) за намалување на димензионалноста на обележјата добиени од

страна на аудио и визуелните дескриптори. Троунг (Truong) ги користи семантичките карактеристики на секој од различните видео жанрови како што се обележјата за боја и движење. Тој во својот пристап како класификациска структура ги користи C4.5 дрвата на одлучување [2]. Овие автори предлагаат решение на проблемот за класификација на видео податоците кои се дел само од највисоката класна хиерархија, односно видео податоци кои можат да се категоризираат во неколку различни основни видео жанрови како што се цртани филмови, спорт, реклами, вести и друго.

Xavier во своето истражување се придвижува едно ниво подолу во класната хиерархија на видео податоците и го разгледува проблемот на класификација на видео податоци со спортска содржина (хокеј, американски фудбал, кошарка и фудбал) [3]. Како класификациски модел користи Hidden Markov Model (HMM), додека видео податоците ги опишува со помош на дескриптори за боја и движење.

Во пристапот понуден од страна на Такаги (S. Takagi) се користење на статистичка анализа на движења на камерата, како што се поставување на кадар, зумирање и вртење, спортските видеа се категоризираат во сумо, тенис, бејзбол, фудбал и американски фудбал [4]. J. Assfiag во своето истражување видео податоците ги дели на слики, а потоа секоја од сликите ја класифицира во соодветната класа терен, играч, публика, коментаторско студио [5]. Индексирање на видео секвенци од жанрот *вести* [6], како и семантичка класификација на кошаркарски секвенци во категориите погодок, фаул или борба за топка [7] се пристапи во кои се користат обележјата базирани на рабови.

Имајќи ги предвид претходните пристапи, во овој труд предлагаме класификациска архитектура за хиерархиска класификација на видео податоци која соодветствува на природната односно вистинската хиерархија на видео податоците во чија содржина преовладуваат пејсажи од природа и градски пејсажи. Како класификатори во оваа архитектура ги користиме машините со носечки вектори. За секој видео податок ги издвоивме неговите клучни слики користејќи го алгоритмот

предложен од страна на Zhonghua Sun [14]. Клучните слики се опишани со обележја што соодветствуваат на неколку различни видови на дескриптори од MPEG-7 стандардот.

Остатокот од овој труд е организиран на следниов начин: Во вториот дел се опишани неколку различни видови дескриптори од MPEG-7 стандардот кои ги користевме во процесот на екстракција на обележја од видео податоците. Деталниот опис на нашата хиерархиска архитектура за класификација на видео податоци е даден во третиот дел, по што следи анализата на експерименталните резултати. Заклучокот и насоките за идните истражувања се дадени во петтиот дел.

2. ЕКСТРАКЦИЈА НА ОБЕЛЕЖЈА

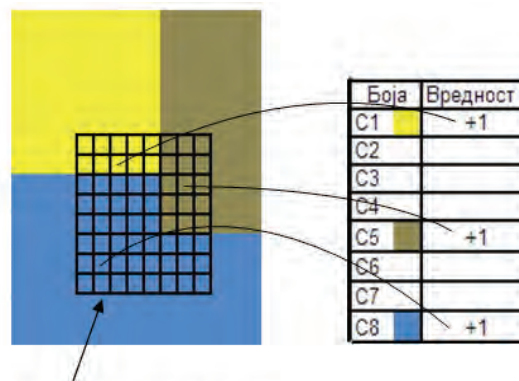
Пред да преминеме на процесот на класификација, потребно е видео податоците да бидат опишани со соодветни дескриптори. Ефикасноста на која било класификациска метода зависи од ефикасноста на содржинската репрезентација на самите податоци добиена со екстракција на нивните обележја. Процесот на екстракција на обележја го започнавме со автоматско одделување на различните кадри кај секој видео податок. Клучните слики кои ги екстрахиравме од добиените кадри со помош на алгоритамот предложен од страна на Zhonghua Sun [14] беа нашиот репрезент на видео податоците. Овие слики дополнително беа опишани со различни видови на MPEG-7 дескриптори базирани на боја и текстура. Multimedia Content Description Interface, генерално познат како MPEG-7, претставува стандард за мултимедијално содржинско опишување за голем број на различни видови податоци меѓу кои слики, видео или аудио податоци. MPEG-7 дефинира три главни категории на визуелни дескриптори за слика: боја, текстура и форма [8][9]. Од искуството во претходни искуства и истражувања направени во областа на обработката на слика [21][22][23] ги одбравме дескрипторите базирани на боја и текстура за опишување на видео податоците, кои во спореба со другите дескриптори даваат подобри резултати.

2.1. Дескриптори базирани на боја

MPEG-7 стандардот дефинира четири групи на дескриптори базирани на боја (доминантна боја, скалабилност на боја, структура на боја и распоред на боја). За опишување на клучните слики од видео примероците ги користевме дескрипторите структура на боја и распоред на боја.

2.2. Структура на боја

Дескрипторот структура на боја е хистограм кој пренесува информација за просторната дистрибуција односно распределбата на боја на ниво на една слика. Тој се извлекува во HMMD просторот со нерамномерна квантизација користејќи 256, 128, 64 или 32 вредности [8][10].



8 x 8 структурен елемент

Слика 2. Креирање на хистограмот кај дескрипторот структура на боја [9]

Пополнувањето на хистограмот на структура на боја е претставен на Слика 2. Да претпоставиме дека постои структурен елемент кој е дел од сликата со големина 8x8. Секој од 64-те пиксели на овој елемент е пополнет со една од трите бои: C₁, C₅ или C₈. Следствено вредностите од хистограмот кои соодветствуваат на овие три бои се зголемуваат за еден штом ќе најдеме до пиксел кој ја содржи соодветната боја.

2.3. Распоред на боја

Дескрипторот *Распоред на боја* нуди компактна репрезентација на бојата присутна во сликата на начин независен од резолуцијата. Обележјата добиени со овој дескриптор се од Y, C_b и C_r просторот на бои [8][11]. Прво сликата се дели на 64 блока за да се избегне зависност од резолуцијата и големината на сликата. За секој од овие блокови се избира единствена боја репрезент за соодветниот блок. Најчесто се избира средната вредност на бојата на сите пиксели опфатени во блокот. Со определување на најрепрезентативната боја за секој од 64-те блокови, се генерира слика со димензии 8x8 пиксели, и секој пиксел на новодобиената слика ја содржи репрезентативната боја на блокот од оргиналната слика што го репрезентира. Следејќи ја 8x8 DCT трансформацијата се издвојуваат 64-те коефициенти.

2.4. MPEG-7 дескриптори на текстура

Во MPEG-7 стандардот дефинирани се три видови на дескриптори на текстура: *хомогеност на текстура*, *пребарување на текстура* и *хистограм на рабови* [8]. Во нашите истражувања беа

користени дескрипторите *хомогеност на текстура* и *хистограм на рабови*.

2.5. Хомогеност на текстура

Дескрипторот *хомогеност на текстура* се состои од 62 коефициенти. Ознаката е

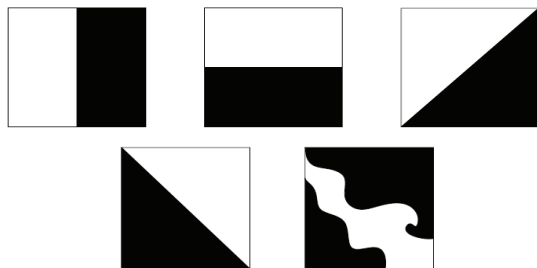
$$HTD = [f_{DC}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}]$$

Првите два атрибути f_{DC} и f_{SD} се средната вредност и стандардната девијација на сликата. Останатите 60 атрибути се репрезенти на средната енергија и енергетската девијација добиени од множеството на фреквентни канали. Сите овие вредности се нелинеарно скалирани и квантизирани [8].

Постојат 30 фреквенциски канали дефинирани во фреквенциското подрачје за дескрипторот *хомогеност на текстура*, добиени со поделба на подрачјето по аголен и радијален правец. Фреквенциското подрачје е рамномерно поделено во аголен правец со интервали од 30 степени, на вкупно 6 региони, додека радијално поделбата е со октална скала на вкупно 5 радиуси. Средната вредност на енергијата и нејзината девијација се пресметуваат за секој од 30-те канали.

2.6. Хистограм на рабови

Хистограмот на рабови ја репрезентира локалната просторна распределба на рабови во сликата [12][13]. Постојат вкупно пет видови на рабови дефинирани во MPEG-7. Тие се претставени на Слика 4.



Слика 4. Пет видови на рабови

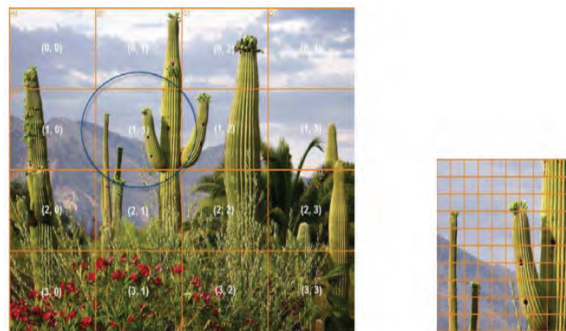
Со делење на сликата на 4x4 потслики како што е прикажано на Слика 5 се добиваат вкупно 16 потслики. Потоа, секоја од овие потслики се дели на непреклопувачки блокови. На Слика 5 дополнително е претставен и процесот на издвојување на блокови од секоја потслика.

За секој регион се прави хистограм со пет вредности кој ја претставува насоката на рабовите во секој од регионите. Хистограмот за целата слика се состои од 80 вредности:

$$H = [h(0,0), h(0,1), \dots, h(1,0), h(1,1), \dots, h(3,3)]$$

Каде $h(i, j)$ е хистограмот на регионот (i, j) ,

$$h(i, j) = [h(0), h(1), h(2), h(3), h(4)] \quad (i, j \in \{0, 1, 2, 3\})$$



Слика 5. Дефиниција на потслика и блок

Секој вектор $h(i, j)$ се состои од пет елементи $h(0)$, $h(1)$, $h(2)$, $h(3)$ и $h(4)$ кои ја претставуваат релативната пополнетост на секој регион со вертикални рабови, хоризонтални рабови, рабови под агол од 45°, рабови под агол од 135° и ненасочени рабови, соодветно.

3. КЛАСНА ХИЕРАРХИЈА И КЛАСИФИКАЦИСКА АРХИТЕКТУРА

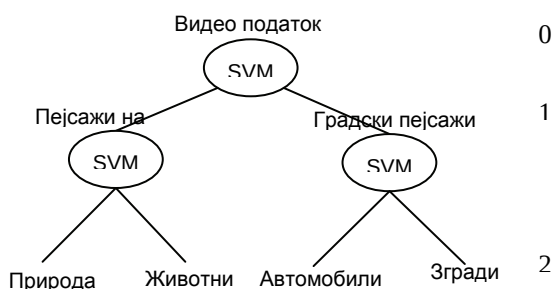
Во овој труд е опишана нова хиерархиска класификациска архитектура, чија хиерархија во целост ќе ја отсликува природната хиерархиска претстава на видео податоците. Природната класна хиерархија ја дефинира содржинската и логичката поврзаност меѓу класите и поткласите дефинирани во различните нивоа од хиерархијата. Имајќи го ваквиот пристап видео податоците може да бидат класифицирани во една или повеќе класи дефинирани во хиерархијата. Една класа може да содржи неколку поткласи, а секоја поткласа може да биде поделена понатаму на неколку поткласи.

Најдобра претстава на природната класна хиерархија може да се постигне со структурата дрво. Секој јазел од дрвото е репрезент на една класа од класната хиерархија. Врската родител-дете во дрво структурата соодветствува на врската класа-поткласа во хиерархијата. Притоа коренот на дрвото претставува репрезент на витуелната класа „Видео податок“, класа која ги обединува сите видео податоци. За оваа класа да се разликува од останатите дефинирани класи претпоставуваме дека таа се наоѓа во коренот на дрвото. Јазлите кои се деца на коренот го дефинираат првото ниво. Нивните деца го дефинираат второто ниво итн. Поради хиерархиската претстава јасно е дека класите во пониските нивоа го опишуваат видео податокот со повеќе детали за разлика од класите кои се дел од повисоките нивоа на класната хиерархија каде видео податокот е опишан со помалку детали.

Вака дефинираната дрво структура ја искористивме за изградба на нашата хиерархиска класификациска архитектура. Таа го решава конкретниот проблем на класификација односно категоризација на видео податоци чија хиерархиска претстава се состои од две нивоа. Во првото ниво од оваа хиерархија се класифицираат

видео податоци чија доминантна содржина се пејсажите на природа и градските пејсажи. Второто ниво на хиерархијата е дефинирано со класите „Животни“ и „Природа“ како репрезенти, односно поткласи на видео податоците чија доминантна содржина се пејсажи од природа, и со класите „Автомобили“ и „Згради“ како репрезенти, односно поткласи на видео податоците чија доминантна содржина се градските пејсажи.

На Слика 6 е прикажана архитектурата дизајнирана за решавање на 4-класен проблем. Класификацијата на видео податоците во нејзините јазли се прави со помош на машини со носечки вектори.



Слика 6. Класификациска архитектура

Фазата на обучување на нултото ниво од архитектурата ја реализиравме со помош на слики чија содржина е базирана на пејсажи на природа и градски пејсажи. Во ова ниво од класната хиерархија за обучување на SVM класификаторот користевме комбинација на обележја од соодветните слики, добиени со екстракција од MPEG-7 дескрипторите *распоред на боја* и *структура на боја*. Од своја страна, сликите базирани на пејсажи на природа се составени од слики чија доминантна содржина се „Животни“ и „Природа“. Овие слики ги користевме во процесот на обучување на левото потстебло од првото ниво во нашата класна архитектура. Тоа значи дека SVM класификаторот кој се наоѓа во левото потстебло на првото ниво од класната хиерархија се обучува со сликите „Животни“ како позитивни примероци и сликите „Природа“ како негативни примероци. Обележјата кои ги користевме за опис на овие слики во процесот на обучување на класификаторот ги добивме со помош на MPEG-7 дескрипторите *распоред на боја*, *структура на боја* и *хомогеност на текстура*. Во процесот на обучување на десното потстебло од првото ниво ги користевме сликите базирани на градски пејсажи кои се составени од слики чија доминантна содржина се „Автомобили“ и „Згради“. Соодветниот SVM класификатор од десното потстебло на првото ниво од класната хиерархија се обучува со сликите „Автомобили“ како позитивни примероци и сликите „Згради“ како негативни примероци. Единствениот дескриптор кој го употребивме за опишување на

овие слики беше дескрипторот хистограм на рабови.

Фазата на тестирање на оваа архитектура се состои од три последователни чекори. Во првиот чекор од секој видео податок се извлекуваат клучните слики со помош на алгоритмот претставен од страна на Zhonghua Sun [14]. Овој чекор е особено значаен, бидејќи само овие слики се користат во понатамошниот процес на класификација како единствени репрезенти на видео податокот. Вториот чекор е екстракција на обележјата од клучните слики. Тие во различните нивоа на класната хиерархија се добиваат на ист начин како и обележјата користени за примероците во процесот на обучување. Третиот чекор се состои од препознавање на клучните слики. Препознавањето на секој од примероците започнува од коренот на дрвото. Во секој од јазлите се донесува одлука за припадноста на примерокот во една од двете класи кои го претставуваат соодветниот јазел. Овој процес се повторува се додека не дојдеме до листовите на дрвото каде впрочем и завршува класификацијата.

4. ЕКСПЕРИМЕНТИ

Податочното множество за обучување на овој систем беше составено од 2000 слики базирани на пејсажи на градови и природа. Класификаторите во понудената архитектура беа тренирани односно обучувани со целото податочно множество предвидено за тренирање.

За да ги евалуираме перформансите на оваа архитектура беа направени експерименти со видео податоци со различна содржина. Вкупното времетраење на видео материјалот беше околу 20 минути. Содржински кај нив преовладуваа пејсажите на природа и градските пејсажи кои од своја страна беа репрезентирани со поголема деталност преку содржините „Животни“ и „Природа“ односно „Згради“ и „Автомобили“. За да ја зголемиме веродостојноста на добиените резултати одбравме тестирачки видео податоци чија вистинска класа на припадност беше утврдена со мануелна визуелна анализа. Поголемиот дел од користените видео секвенци беа со послаб квалитет и ниска резолуција што беше дополнителен предизвик при процесот на класификација.

За да ги евалуираме перформансите на хиерархиската архитектура, експериментите кои ги направивме се однесуваа на споредба на различните дескриптори со кои беа опишани видео податоците. Во согласност со ова експериментите ги групиравме во 4 различни групи според множествата на обележја кои беа користени за опишување на сликите и клучните слики од видео податоците при процесот на обучување односно тестирање. Во првата група на експерименти го користевме множеството обележја добиено од MPEG-7 дескрипторот *распределба на боја*. Во втората група на

експерименти податоците беа опишани со дескрипторот *структура на боја*. Третата и четвртата група на експерименти беа направени со множествата обележја добиени од дескрипторите *хомогеност на текстура* и *хистограм на рабови* соодветно.

Во Табела 1 се претставени процентите на погрешно класифицирани примероци за секоја од класите поделно како и процентите на погрешна класификација во првото и второто ниво на класификациската архитектура. Процентот на грешка во второто ниво всушност претставува и грешка на системот во целина. Од добиените резултати може да се забележи дека за грешката во првото ниво на системот се добиени споредбени резултати во Е. 1, Е. 2, Е. 3 и Е. 4 групите на експерименти, додека во второто ниво таа варира од 50.3% за групата на експерименти направена само со дескрипторот хомогеност на текстура (Е. 4), до 24.9% за групата на експерименти направена со дескрипторот распределба на боја (Е. 1). Од анализата на направените експерименти и резултатите кои ги добивме дојдовме до заклучок дека дескрипторите базирани на текстура не се многу погодни за опишување на видео податоци чија тематика се пејсажите на природа, додека од друга страна тие одлично ги репрезентираат видео податоците чија тематика се градските пејсажи. За дескрипторите базирани на боја дојдовме до заклучок дека подеднакво успешно ги опишуваат видео податоците чија тематика се пејсажите на природа или градските пејсажи.

Врз основа на добиените резултати од овие експерименти направивме уште една, петта група

на експерименти (Е. 5) во која користевме комбинација на различни видови на дескриптори за опишување на видео податоците. Комбинацијата на дескриптори која ја направивме во оваа група на експерименти се темели на предложената архитектура од претходното поглавје. За опишување на видео податоците користевме комбинација од дескрипторите за боја за опишување на податоците од обучувачките и тестирачките множества во првото ниво на класната хиерархија. Во второто ниво од хиерархијата користевме комбинација од дескрипторите на боја и дескрипторот *хомогеност на текстура* за податоците базирани на пејсажи на природа. За податоците базирани на градски пејсажи од второто ниво на класната хиерархија го користевме дескрипторот хистограм на рабови. Процентот на грешка за оваа група на експерименти изнесува 17.7%. Преку анализа дојдовме до заклучок дека оваа комбинација на дескриптори даваат подобри перформанси во просек од 10% во второто ниво на хиерархија во однос на резултатите кои ги постигнавме со употреба на само еден вид на дескриптор. Во првото ниво на хиерархија предноста не е толку очигледна и таа се движи од 0.6% до 11.9%. Притоа можеме да кажеме дека со многу мало зголемување на комплексноста на класификациската метода што произлегува од зголемениот број на обележја употребени за опишување на видео податоците, добивме значителен процент на подобрување во однос на прецизноста.

	Е. 1	Е. 2	Е. 3	Е. 4	Е. 5
Грешка на прво ниво	10.3	7.2	18.5	8.7	6.6
Природа	11.5	8.5	18.4	8.4	6.9
Животни	24.7	34.2	34.8	95.0	24.7
Згради	28.1	35.1	33.3	28.0	28.0
Автомобили	91.6	41.6	4.2	0.0	0.0
Грешка на второ ниво	24.9	26.2	27.2	50.3	17.7

Табела 1. Процентот на грешка на класификационата архитектура

5. ЗАКЛУЧОК

Класификацијата на видео податоците е многу важен сегмент во содржински базираниот видео пристап за да имаме поголема ефикасност и брзина при нивното пребарување. Поради тоа во овој труд е претставена една архитектура за класификација на видео податоци. При класификацијата се употребува комбинација од обележјата на видео секвенците екстрахирани со помош на MPEG-7 дескрипторите за боја и текстура. Архитектурата која ја предлагаме е хиерархиска, базирана на SVM класификатори.

Употребено е бинарно класификациско дрво кое ја отсликува природната класна хиерархија на видео податоците. Архитектурата е дизајнирана со цел да се постигнат супериорни перформанси во однос на брзината и прецизноста на класификација на токму овој вид на податоци.

ЛИТЕРАТУРА

- [1] L.-Q. Xu and Y. Li, .Video classification using spatialtemporal features and PCA,. in *Int. Conf. Multimedia and Expo*, (Baltimore, MD, USA), pp. 345-348, July 6-9, 2003.

- [2] B. T. Truong, S. Venkatesh, and C. Dorai, .Automatic genre identification for content-based video categorization, in *Proc. of Int. conf. Pattern Recognition*, (Barcelona, Spain), vol. 4, pp. 230-233, Sep. 3-8, 2000.
- [3] Gibert, H. Li, and D. Doermann, .Sports video categorizing using hmm, in *Int. Conf. Multimedia and Expo*, (Baltimore, MD, USA), pp. 345-348, July 6-9, 2003.
- [4] S. Takagi, S. Hattori, K. Yokoyama, A. Kodate, and H. Tominaga, Sports video categorizing method using camera motion parameters, in *Int. Conf. Multimedia and Expo*, (Baltimore, MD, USA), pp. 461-464, July 6-9, 2003.
- [5] J. Assfiag, M. Bertini, C. Colombo, and A. D. Bimbo, .Semantic annotation of sports videos, *IEEE Multimedia*, vol. 9, pp. 52-60, June 2002.
- [6] V. Suresh, S. Palanivel, and B. Yegnanarayana, .Anchorperson indexing and visual-table-of-contents generation for tv news, in *Proc. Workshop on Biometric Challenges Arising out of Theory and Practice*, (Cambridge, UK), pp. 59-62, Aug 22-25, 2004.
- [7] M. H. Lee, S. Nepal, and U. Srinivasan, .Edge-based semantic classification of sports video sequences, in *Int. Conf. Multimedia and Expo*, (Baltimore, MD, USA), pp. 157-160, July 6-9, 2003.
- [8] B.S. Manjunath (Editor), Philippe Salembier (Editor), and Thomas Sikora (Editor): Introduction to MPEG-7: Multimedia Content Description Interface. Wiley & Sons, April 2002
- [9] M.-. V. Group, "Text of iso/iec 15938-3/fdis information technology – multimedia content description interface – part 3 visual," in ISO/IEC/JTC1/SC29/WG11/N4358, Sydney, 2001.
- [10] Manjunath BS, Ohm J-R, Vasudevan VV and Yamada A (2001) Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology* 11:703-715.
- [11] L. Cieplinski, "MPEG-7 color descriptors and their applications, ", in *Lecture Notes in Computer Science: Computer Analysis of Images and patterns*, Springer 2001, p11
- [12] D. K. Park, Y. S. Jeon, and C. S. Won, Efficient use of local edge histogram descriptor, *ser Proceedings of the 2000 ACM workshops on Multimedia*. Los Angeles, California, United States: ACM Press, 2000.
- [13] J. Wu, X. Zhang, and J. Zhou, "Vehicle detection in static road images with pca-and-wavelet-based classifier," in *Proceedings of IEEE Intelligent Transportation Systems*, vol. August, Oakland (CA), USA, 2001.
- [14] Zhonghua Sun, Kebin Jia, Hexin Chen. "Video Key Frame Extraction Based on Spatial-temporal Color Distribution", in *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp:196-199
- [15] Torch: a modular machine learning software library. Technical Report IDIAP-RR 02-46, IDIAP, 200
- [16] Gjorgji Madzarov, Dejan Gjorgjevikj, Ivica Dimitrovski, "A MULTI-CLASS SVM CLASSIFIER UTILIZING BINARY DECISION TREE", 11th International Multiconference INFORMATION SOCIETY-IS 2008 Ljubljana, Proceedings Volume-A, ISSN 1581-9973, pp.77-81, 13-17 October, Slovenia, 2008

Summary

HIERARCHICAL VIDEO CLASSIFICATION

Gjorgji Madzarov, Ivica Dimitrovski, Suzana Loskovska, Dejan Gjorgjevikj

"Ss. Cyril and Methodius" University – Skopje, Faculty of Electrical Engineering and Information Technologies – Skopje, Karpoš II bb, PO Box 574, 1000 Skopje,

Abstract - In this study we propose a novel architecture of Support Vector Machines (SVM) classifiers for classifying video documents in classes organized in a predefined class hierarchy. This architecture of SVM classifiers was designed to provide superior recognition speed and accuracy utilizing a binary classification tree. It solves a four-class video classification problem, having a two level hierarchy. The classification data consisted of automatically detected key frames, which were described by the MPEG-7 standard. The experimental results show that the classification architecture achieves high classification accuracy, especially in the first level of the hierarchy where the classification was made between different genres of videos.

Keywords – Support Vector Machines (SVM), MPEG-7 descriptors, hierarchical classification.