

**Нов български университет
департамент “Телекомуникации”**

инж. Георги Костадинов Петров

**Многомерна цифрова обработка и анализ на
последователности от изображения**

АВТОРЕФЕРАТ

на дисертация за получаване на научно образователна степен

“ДОКТОР НА НБУ”

по научна специалност: 02.21.07 **“Автоматизирани системи за обработка
на информация и управление (по отрасли)”**

София © 2008

**Нов български университет
департамент “Телекомуникации”**

инж. Георги Костадинов Петров

**Многомерна цифрова обработка и анализ на
последователности от изображения**

АВТОРЕФЕРАТ

На дисертация за получаване на научно образователна степен

“ДОКТОР НА НБУ”

по научна специалност: 02.21.07 **“Автоматизирани системи за обработка
на информация и управление (по отрасли)”**

Научен ръководител:
доц. д-р Панайот Илиев

Рецензенти:
проф. Филип Филипов, д.н.
доц. д-р Пламен Цветков

София © 2008

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност на проблема

Интернет и мобилните приложения са основният фактор, влияещ върху масовизацията на съвременните мултимедийни технологии. Развитие получават цифровото телевизионно и видео разпръскване, видеото по заявка, мобилната телевизия, системите за контрол и наблюдение, медицина и други. Тези приложения изискват наличието на съвременни, бързи и мащабируеми хардуерни и софтуерни архитектури за обработка, анализ, индексирание, компресия и предаване на информацията. Ползваните методи за обработка трябва да осигурят високо бързодействие и достоверност на съхраняваната информация.

Цели и задачи на дисертацията

Целта на този дисертационен труд е модификация, разработка и изследване на функционален модел за многомерна вероятностно статистическа обработка и анализ на последователности от изображения. За реализацията на това са формулирани следните основни цели:

1. Да се избере и софтуерно реализира метод за предварителна филтрация на последователности от изображения.
2. Модифициране на метод за откриване на движение, ползван в системите за видео наблюдение, така че да може се прилага при анализа на видео филми.
3. Модифициране на 2D и 3D ентропийни модели за реализиране на многоканална ентропийна сегментация.
4. Избор на основните параметри на обектите и сегментираните последователности от изображения, които да се ползват при тяхната индексация.
5. Разработване на метод за блокова обработка и анализ на последователности от изображения чрез 4D хистограми.
6. Разработване на алгоритъм за определяне на грешката при автоматизираното фрагментиране на видео филми, и да се избере методика за обективна метрика на качеството.
7. Разработване на софтуерни модули за потвърждаване на получените експериментални резултати на разработените методи за многомерна обработка на последователности от изображения.

Структура на дисертационния труд

1. АНАЛИЗ НА СЪВРЕМЕННОТО СЪСТОЯНИЕ ПО ТЕМАТА

Методи за предварителна обработка и филтрация на последователности от изображения

Методи за сегментация на движението в последователности от изображения

Методи за обектна сегментация и векторен анализ на последователности от изображения

ЦЕЛИ И ЗАДАЧИ НА ДИСЕРТАЦИОННИЯ ТРУД

2. ИЗБОР И МОДИФИЦИРАНЕ И НА МЕТОДИ ЗА ПРЕДВАРИТЕЛНА ВЕРОЯТНОСТНО СТАТИСТИЧЕСКА ОБРАБОТКА НА ПОСЛЕДОВАТЕЛНОСТИ ОТ ДВИЖЕЩИ СЕ ИЗОБРАЖЕНИЯ

Разработка и изследване на метод за 3D хистограмно групиране на последователности от изображения, базиран върху метода за 2D хистограмно групиране

Изследване и избор на едно кадров и темпорален модов или медианен адаптивен филтър за филтрация на изображения, определяне на основните изисквания

3. МОДИФИЦИРАНЕ И ПОДОБРЯВАНЕ НА МЕТОД ЗА ОТКРИВАНЕ НА ДВИЖЕНИЕ, ИЗПОЛЗВАН В СИСТЕМИ ЗА ВИДЕО ОХРАНА, С ЦЕЛ ПРИЛАГАНЕ В СИСТЕМИ ЗА ФРАГМЕНТИРАНЕ НА ЦИФРОВИ ВИДЕО ФИЛМИ

Модификация на 2D хистограмен метод за детекция на движение чрез критериите на Колмогоров, Пийрсън и комплексен критерий чрез ползване на 3D хистограми

Анализ на многокритериални не адаптивни методи за детекция на движение

Разработка на метод за детекция на глобални и локални ключови кадри в последователности от изображения с отчитане на цветната мода в два последователни кадъра

4. МОДИФИЦИРАНЕ НА АЛГОРИТМИ ЗА 2D И 3D ХИСТОГРАМНА ЕНТРОПИЙНА СЕГМЕНТАЦИЯ НА СИВИ ПОЛУТОНОВИ ИЗОБРАЖЕНИЯ

Метод за многоканална 2D и 3D хистограмна и ентропийна (градиентна) сегментация на изображения

Избор на ключови параметри за обектния многомерен векторен анализ на последователности от ентропийно сегментирани изображения и параметрична филтрация

5. РАЗРАБОТКА НА ФУНКЦИОНАЛЕН МОДЕЛ ЗА БЛОКОВА 4D ХИСТОГРАМНА ОБРАБОТКА НА СЕМАНТИЧНО СВЪРЗАНИ ПОСЛЕДОВАТЕЛНОСТИ ОТ ИЗОБРАЖЕНИЯ

Разработка на 4D хистограмен модел за обработка на визуални сцени

Дефиниране на референтен фонен кадър за детекция на подвижните обекти

6. МЕТОДИКА ЗА ОТКРИВАНЕ НА ГРЕШКАТА ПРИ АВТОМАТИЗИРАНО ДЕТЕКТИРАНЕ НА КЛЮЧОВИ КАДРИ И МЕТРИКА НА КАЧЕСТВОТО НА ЦИФРОВИТЕ ИЗОБРАЖЕНИЯ

Дефиниране на TKFC, MKFI и KFDI

Оценка на PSNR и MSE за метрика на качеството

Дефиниране на локалния контраст в ентропийно сегментираните

7. СОФТУЕРНИ МОДУЛИ И ПРИЛОЖЕНИЯ ЗА АНАЛИЗ НА ПРЕДЛОЖЕНИТЕ В ДИСЕРТАЦИЯТА МЕТОДИ ЗА МНОГОМЕРНА ОБРАБОТКА И АНАЛИЗ

хистограмно хомогенизиране, медианна филтрация, многоканална ентропийна и обектна сегментация, параметрична филтрация, анализ и сравняване на движението в 4D хистограми на различни последователности от изображения

Методологична основа

За нуждите на изследванията са създадени специализирани авторски софтуерни решения, написани на C# и Visual C++. Проектирани са софтуерни филтри за време-пространствена вероятно статистическа филтрация на последователности от изображения, многоканална ентропийна сегментация и параметрична филтрация на обектите. За анализ на експерименталните данни са използвани и други софтуерни продукти. Под формата на приложения са дадени експериментални резултати.

Практическа стойност на разработката

Фокусът на представените в дисертацията вероятно статистически функционални модели, е върху възможността да бъдат подобрени и модифицирани методи за детекция и анализ на движението, използвани в системите за видео охрана и наблюдение и в системите за фрагментиране и индексирание на видео филми. Приложението на предложените в дисертацията методи би имало положителен ефект върху подобряване на производителността, ефективността и автоматизацията на дейностите по индексирание на видео и мултимедийната информация. Информацията от индексиранияте видео записи може да се използва в системите за приоритизиране и планиране на трафика на телекомуникационните доставчици. Предложените методи за автоматизирано сегментиране на видео потоците позволява създаването на автоматизирани системи за класификация на големи обеми от видео бази данни, каквито са: <http://www.youtube.com/>, <http://www.vbox7.com/>.

Публикуване на резултатите от изследването

Основните резултати от изследванията в дисертационния труд са представяни в общо 10 научни публикации по дисертационната тема в сборници с доклади от национални научни симпозиуми с международно участие (A3, A4, A6, A7, A8) и конференции с международно участие (IEEE) (A5-Истанбул, A9, A10-София, WSEAS-A11 София), както и две в украинското списание Computing International Journal (A1, A2).

Структура и обем на дисертацията

Дисертационният труд съдържа специално подбрани за случая: 9 глави (с общо 50 раздела), 100 фигури, 100 таблици, 50 графики и приложения, като общият обем на изложението и приложенията е 152 страници. Приложени са: списък на използваните съкращения и термини, библиография по темата (153 източника), както и приложения с експериментални резултати.

В основата на предложения комплексен функционален модел за многомерна вероятно статистическа цифрова обработка и анализ на последователности от изображения стоят оригинални методи и софтуерни приложения за откриване на движение и сегментиране на последователности от изображения. В предлаганите методи са залегнали математически зависимости, използвани при разработването на софтуерните модули, които реализират функционалния модел за обработка и анализ на последователности от изображения: откриване на движение, филтрация и ентропийна сегментация, използвана за квантуване на изходните изображения на две и повече нива. Предложени са алгоритми за анализ на последователности от изображения, чрез 3D и 4D хистограми и ентропийните им функции. Детайлно са обсъдени статистическите и функционални модели за откриване на движение, като е отделено особено внимание на методите за финализиране на процеса по детектиране на ключови кадри, използвайки статистическите критерии на Колмогоров, Пиърсън и Комплексен. Предложени са процедури за детекция на глобални и локални ключови кадри, които позволяват автоматизираното разделяне на видео филми на семантично свързани визуални сцени.

Приложени са комплексни процедури за откриване и проследяване на подвижни обекти, класификация и параметрична филтрация. Предложен е нов модел за анализ на визуални сцени, чрез техните 4D нормирани хистограми и ентропийни функции, подходящ за сравняването на групи от визуални сцени. В този раздел е предложен и имплементиран 3D филтър на модата и 3D медианен филтър за обработка на последователности от изображения. Анализът на последователностите от изображения се свежда до анализ на многомерното векторно пространство на сегментираните визуални сцени. Описани са основните възможности на разработените софтуерни модули реализиращи представените методи на обработка. В приложения са представени експерименталните резултати.

Благодарности

Издавам благодарност към научния ми Ръководител доц. д-р Панайот Илиев и към рецензентите си, както и на екипа от специалисти и преподаватели към департамент “Телекомуникации” в НБУ за представената възможност за развитие.

Благодаря на своите родители.

ОСНОВНО СЪДЪРЖАНИЕ

НА ДИСЕРТАЦИОННИЯ ТРУД

Глава 1

АНАЛИЗ НА СЪВРЕМЕННОТО СЪСТОЯНИЕ ПО ТЕМАТА

В тази глава са разгледани методите използвани за многомерна цифрова обработка и анализ на последователности от движещи се изображения. Настоящата дисертация акцентира върху модификацията, усъвършенстването и комбинирането на известни вероятно статистически функционални модели за сегментиране на движението в системите за видео охрана и наблюдение (P. Iliev and L. Tsekov, 1993).

Хипотезата която изказваме е да бъдат модифицирани методите за обработка и детекция на целите от системи за видео охрана, така че да станат приложими в системите за анализ, фрагментация и индексирание на видео филми. Анализирайки литературните източници (Greg P., Ramin Z., 2005) е изказано предположението, че комбинирането на критериите на Колмогоров, Пийрсън и комплексен критерии за анализ на 3D хистограмите ще доведе до значително повишаване точността на детекцията на движение в сравнение с 2D хистограмния модел. Отчитайки особеностите на видео филмите е предложена модификация на цветния хистограмен подход (Yu-Jin Zhang, p41-57, 2007) и контраста на изображенията (Stefan Winkler, p71-103, 2005), като елемент на психо-визуалния модел за сегментиране на визуални сцени.

Основна разлика на предлагания време пространствен 2D ентпориен модел даден от (Yu-Fei Ma, Hong-Jiang Z. 2006) е използването вероятно статистически критерии за 3D хистограмите. Друга разлика с (Roland M. 2005) е ползването на глобални вероятно статистически методи.

Анализирани са вероятно статистическите алгоритми за филтрация на видео филми (A2, A8, A9, 22, 32, 50, 76, 83, 84, 123, 136, 145), както и други методи за предварителна обработка чрез бърза Фурие и косинусова, Уолш-Адамар и Уеивлет трансформация (2, 11, 14, 33, 52, 53, 59, 60, 65, 72, 92), излизаци извън рамките на избрания от нас подход. За предварителна филтрация са изследвани два типа филтри: адаптивен модов (Д.Дончев, 2007) и медианен (Turney, R.D.;

Reza, A.M.; Delva, J.G.R, 1999). Избрано е да бъде модифициран медианен филтър, чрез подбор на тегловни коефициенти (Gonzalo R. Arce, p139-180, 2005), което го прави подходящ за софтуерна реализация. Като основен информационен показател е избрано да бъде ползвана 2D и 3D ентропийната функция (С.К.Leung, F.К.Lam, 1997), която се прилага многократно, за да се осигури квантуване на изображенията на 4 или 8 нива. На база анализа на 3D хистограмния модел за обработка и анализ на изображения е разработена хипотеза за изпълването на 4D нормирани хистограми в цялата последователност от семантично свързани изображения.

Избрани са метрики за класификация на хистограмите на изображенията (Pavel Z., Giuseppe A., Vlastislav D.,2005), като е предложена комбинирана метрика за качеството на изображенията и точността при фрагментация с ключови кадри (Heinrich H.B., Seong-Whan Lee, Tomaso A.P., C.W. (Eds.), p230-240, 2002).

Изводи към Глава 1

Модификацията на разгледаните методи и създаването на нови вероятно статистически функционални модели и софтуерни приложения следва да бъдат съобразено с изискването за приложимост в системите за фрагментация и индексирание на видео филми.

Приноси към Глава 1

На база извършения литературен обзор по темата са представени основните хипотези, реализацията на които цели адаптация и подобряване на известни вероятно-статистически методи за обработка, детекция на движение, сегментиране и анализ на последователности от изображения.

Глава 2

ИЗБОР И МОДИФИЦИРАНЕ И НА МЕТОДИ ЗА ПРЕДВАРИТЕЛНА ВЕРОЯТНОСТНО СТАТИСТИЧЕСКА ОБРАБОТКА НА ПОСЛЕДОВАТЕЛНОСТИ ОТ ДВИЖЕЩИ СЕ ИЗОБРАЖЕНИЯ

Обработка на изображения чрез групиране на 3D хистограмата им

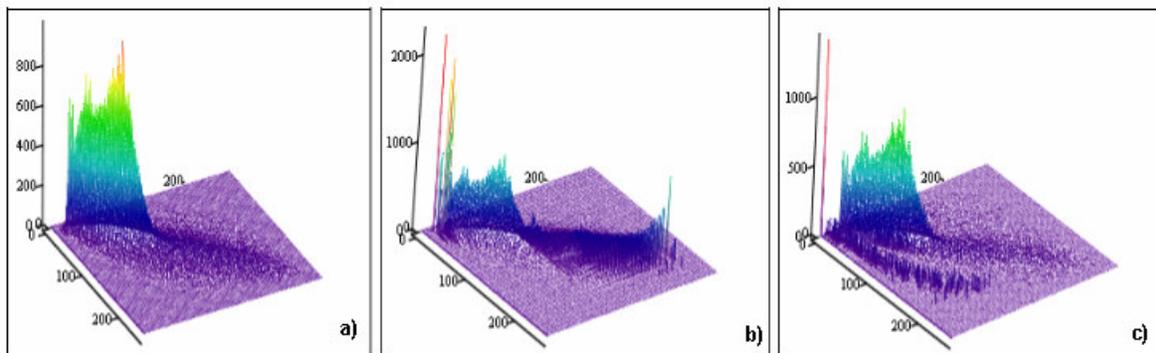
При обработката на видео филми се ползват методи позволяващи бърза обработка. След извършения литературен обзор, бе установено, че масово прилаганите и подходящи за софтуерна реализация методи за обработка на изображения са чрез манипулации в техните 2D

хистограми: групиране, усилване и други целящи промяна на контраста операции. Тъй като процедурите за филтрация на изображения с ядра с размери над 5x5 пиксела отнемат много време при софтуерна реализация, резонно е да се търсят други еквивалентни алгоритми, ускоряващи процесът на филтрация на големи изображения (А6). Описаните в Глава 1 методи за хистограмно групиране в 2D хистограмата, може да бъдат адаптирани в 3D хистограмата на изображенията (фиг. 2.6).



Фиг. 2.6. Оригинално изображение – a), групиране в 3D хистограма – b), разлика в изображенията – c)

Групирането в 3D хистограмата позволява оптимално сливане на съседни двойки пиксели. Групирането в 2D хистограмата прави това, разглеждайки вероятностното разпределение на отделните пиксели самостоятелно (фиг. 2.8). При групиране на 3D хистограмите на прекалено тъмни не контрастни изображения е възможно образуването на дефекти изразяващи се в тъмни пиксели. Предложеният метод позволява хомогенизирането на отделни области в изображенията без съществено да нарушава параметрите на обектите.



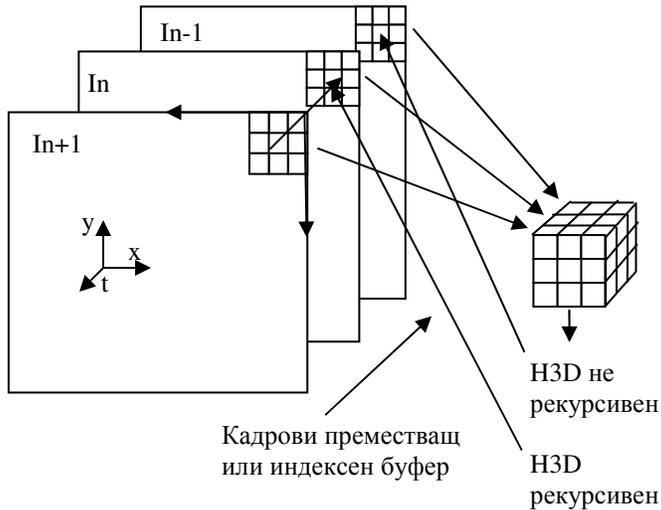
Фиг. 2.8 Визуализация на 3D хистограма на оригиналното изображения – a), 3D хистограма на обработено изображение чрез групиране в неговата 2D хистограма – b), и групирана 3D хистограма – c)

Адаптивна статистическа филтрация

С цел хомогенизиране на отделните изображения в сегментирани визуални сцени са използвани процедури за 2D пространствена и 3D

време пространствена филтрация използваща филтър на модата и медианата. Това осигурява премахването на импулсия шум и хомогенизации на последователностите от изображения.

Предварителната статистическа филтрация е основна стъпка преди ентропийната сегментация на изображения в рамките на една визуална сцена. Така се постига редуциране на случайните величини подлежащи на обработка. Процедурата по 3D филтрацията е дадена на (фиг. 2.10), този тип обработка изисква значително бързодействие на системата в сравнение с 2D филтрите.



Фиг. 2.10. Реализация на не рекурсивен 3D модов/медианен филтър. $In+1$, In , $In-1$ – са поредните изображения.

Сравнени са резултатите, постигнати след ентропийна сегментация на изображения, филтрирани с 3D модов и медианен филтър, както и резултатите, постигнати с 3D рекурсивен медианен филтър с тегловни коефициенти (фиг. 2.12). Обработката на последователности от изображения с 3D филтри е разпространена при компресията с 3D дискретна косинусова трансформация (Westwater R., Furht B. 1997), при едновременно се компресират последователност от 3 до 5 поредни кадъра в блок.



Фиг. 2.12 Ентропийна маска на изображение филтрирано с 3D модов филтър – a), и 3D медианен филтър – b), 2D медианен филтър с тегловни коефициенти - c)

Адаптивността на филтъра се осигурява чрез следене на зоната на рекурсивност, в която попада филтърното ядро спрямо статистически

критерии изчисления на база глобалните 2D хистограми:

математическо очакване:

$$m_l = \sum_{r=0}^k r p_{lr} \quad (1)$$

където: m_l е математическото очакване, а r е стойността на случайната яркост в даден стълб, имащ вероятност p_{lr}

дисперсия:

$$D_l = \sum_{r=0}^k (r - m_l)^2 \cdot p_{lr} \quad (2)$$

коэффициент на отклонение:

$$CV = \frac{D_l}{m_l} \times 100\% \quad (3)$$

където: D_l е дисперсията, а m_l – математическото очакване (средното значение на случайната величина).

зона на рекурсивност:

$$z_l = m_l \pm \frac{D_l}{2} \quad (4)$$

Изводи към Глава 2

- Прилагането на процедури за предварителна филтрация и хомогенизиране на изображенията е основна предпоставка за оптималното им сегментиране.

Приноси към Глава 2

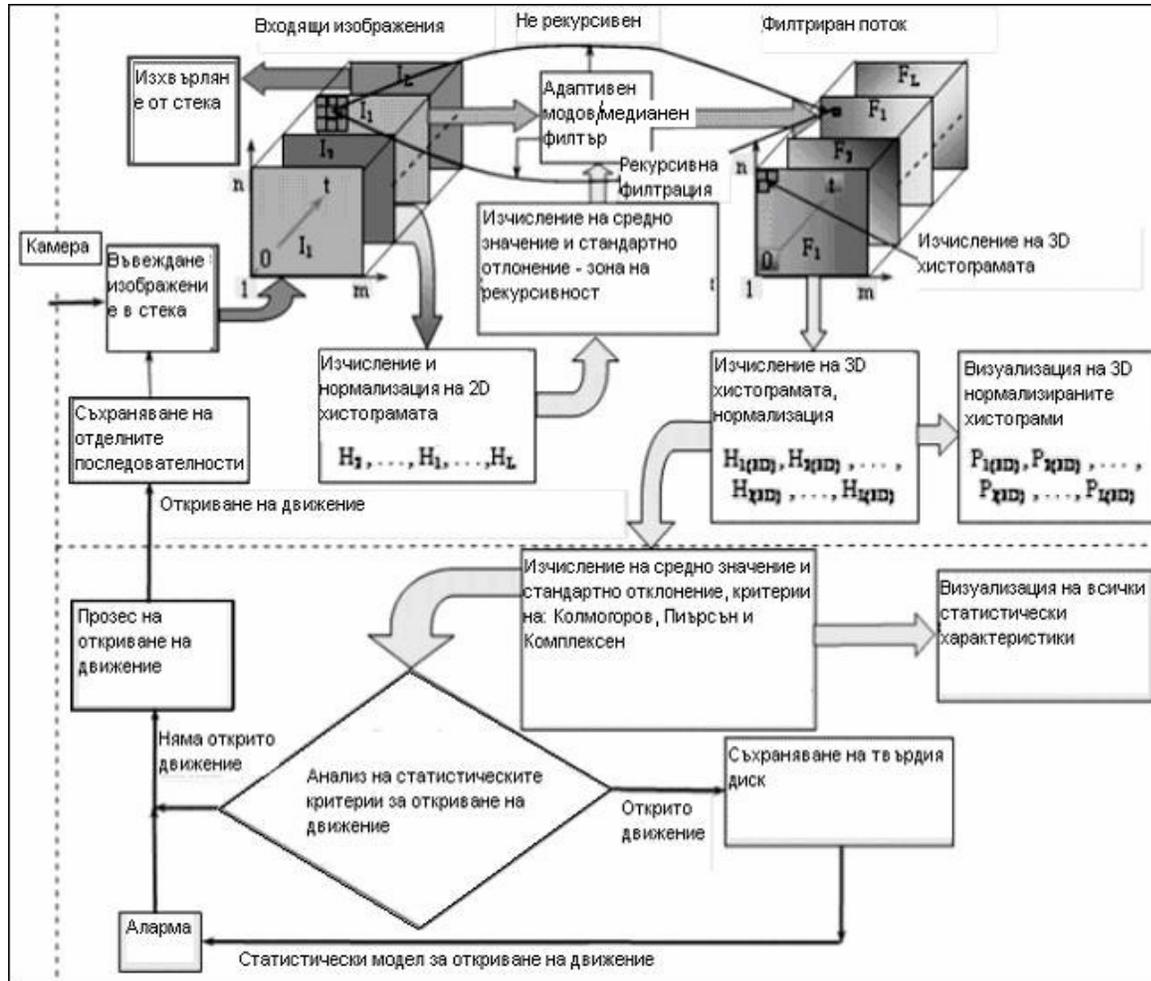
- Модифициран е метода за хистограмно групиране на 2D хистограмите на изображенията, така че да може да се прилага за групиране в 3D хистограмите.
- Опитно е доказано, че адаптивния 3D медианен филтър с тегловни коефициенти, позволява най-висока бързина и качество при хомогенизация на изображенията.

Глава 3

МОДИФИЦИРАНЕ И ПОДОБРЯВАНЕ НА МЕТОД ЗА ОТКРИВАНЕ НА ДВИЖЕНИЕ, ИЗПОЛЗВАН В СИСТЕМИ ЗА ВИДЕО ОХРАНА, С ЦЕЛ ПРИЛАГАНЕ В СИСТЕМИ ЗА ФРАГМЕНТИРАНЕ НА ЦИФРОВИ ВИДЕО ФИЛМИ

В основата на предложения комплексен модел за многомерна цифрова обработка и анализ на последователности от изображения стои модификация на вероятно статистически функционален модел за

откриване на движение (P. Iliev and L. Tsekov, 1993). Постигнато е повишаване чувствителността на детекцията на движение (фиг. 3.1) при слаби изменения, и също така е разработена методика за сегментация на видео филми на база детектирането на глобални и локални ключови кадри (A9). При детекцията на движение се използват полутонни сиви изображения или ентропийните им маски, като се анализират техните 3D хистограми.



Фиг. 3.1 Функционален модел за откриването на движение

Процесът на откриване на движение се осъществява по следните стъпки:

- Прихващане на входящите изображения;
- Адаптивна медианна филтрация;
- Изчисление и нормиране на глобалната 2D статистическа функция на яркостните преходи за всяко едно изображение;

$$f_l(x, y) = \begin{cases} 0 & \text{for } x, y < 0, \\ \sum_{r=0}^k \sum_{r=0}^k p_{l_{rx,ty}} & \text{for } x, y \in [0, k], \\ 1 & \text{for } x, y > k. \end{cases} \quad (5)$$

- Изчисление на статистически средното значение и дисперсията на 3D хистограма;

$$m_{l(2D)} = \sum_{r=0}^k \sum_{r=0}^k \sqrt{r_x^2 + r_y^2} \cdot p_{l(r_x, r_y)} \quad (6)$$

където: $m_{l(2D)}$ е средното значение на 2D хистограмата, r_x и r_y са стойности на яркостта в 3D нормираната хистограма, $p_{l(r_x, r_y)}$ е вероятността в съответната точка, за $r_x \wedge r_y$ $p_{l(r_x, r_y)}$ е вероятността в съответната точка, за $r_x \wedge r_y$.

$$D_{l(2D)} = \sum_{r=0}^k \sum_{r=0}^k [\sqrt{r_x^2 + r_y^2} - m_{l(2D)}]^2 \cdot p_{l(r_x, r_y)} \quad (7)$$

където: $D_{l(2D)}$ е стандартното отклонение.

- Изчисление на статистическите параметри и критерии използвани за анализ на 3D хистограмите (A2):

Критерий на Колмогоров:

$$K = \max_{x,y} f_a(x, y) - f_b(x, y) \quad (8)$$

където: $f_a(x, y)$ и $f_b(x, y)$ са функциите на вероятностното разпределение $P_{l(3D)}^a$ и $P_{l+1(3D)}^b$ за две последователно заснети предварително филтрирани изображения F_1^a и F_{l+1}^b , съответно; $l = 1, 2, \dots, L$.

Критерий на Пиърсън:

$$\chi^2 = (m \times n) \cdot \sum_{r=0}^k \sum_{r=0}^k \frac{(p_b(r_x r_y) - p_a(r_x r_y))^2}{p_a(r_x r_y)} \quad (9)$$

където: $p_{b(rr)}$ и $p_{a(rr)}$ са вероятностите на rr -th сиви-нива на две съседни филтрирани изображения F_{l+1}^b и F_1^a , съответно.

Комплексен критерий (тест за средно значение):

$$Z = \frac{m_a - m_b}{\sqrt{(D_a + D_b)/(m \times n)}} \quad (10)$$

където: $m_{l(2D)}$ и $m_{l+1(2D)}$ са статистически средните значения на две съседни изображения F_1^a и F_{l+1}^b , $D_{l(2D)}^a$ и $D_{l+1(2D)}^b$.

Детектиране на глобални и локални ключови кадри

Поради факта, че ползваме 3 критерия за откриване на движение в последователности от изображения, следва да се дефинират отделни прагове за всеки един от тях (A2). Този подход не е адаптивен, но е достатъчно надежден в системите за охрана и наблюдение, като се прилага по указания по-долу начин. Величините R_K , R_{χ^2} , R_Z подпомагат вземането на решения съгласно следните правила:

$$R_K = \begin{cases} 0, & \text{if } K > t_K \\ 1, & \text{if } K \leq t_K, \end{cases} \quad (11)$$

$$R_{\chi^2} = \begin{cases} 0, & \text{if } \chi^2 > t_{\chi^2} \\ 1, & \text{if } \chi^2 \leq t_{\chi^2}, \end{cases} \quad (12)$$

$$R_Z = \begin{cases} 0, & \text{if } Z > t_Z \\ 1, & \text{if } Z \leq t_Z, \end{cases} \quad (13)$$

където: t_K , t_{χ^2} , t_Z са експериментално определени стойности за отделните статистически критерии K , χ , Z

Крайното решение за наличието на подвижен обект се взема посредством следното правило:

$$R^* = R_K \wedge R_{\chi^2} \wedge R_Z = \begin{cases} 0, & \text{Няма движения} \\ 1, & \text{Има движещ се обект} \end{cases} \quad (14)$$

Проведените експерименти доказват, че изпълнението на това условие указва еднозначно позициите на кадрите, за които сме сигурни, че е настъпила съществена промяна по отношение на предходния.

$$R_i^* = \begin{cases} 1, & ((\Delta K_i \& \Delta \chi_i) > 0) \& (\Delta Z_i < 0) \\ 0, & \text{неспазено} \end{cases} \quad (15)$$

където: i – е поредния номер на кадър в разглежданата последователност, K , χ и Z са статистическите критерии на движението, R_i^* индексен масив

$$M_i^* = \begin{cases} K_i + |Z_i| \text{ за } R_i^* = 1 \\ 0 \text{ за } R_i^* = 0 \end{cases} \quad (16)$$

където: M_i^* е индексен едномерен масив

При семантично фрагментиране на видео филми се налага използването на коригиращ фактор отчитащ модата на цветната хистограма на съседните изображения за които имаме детектиран

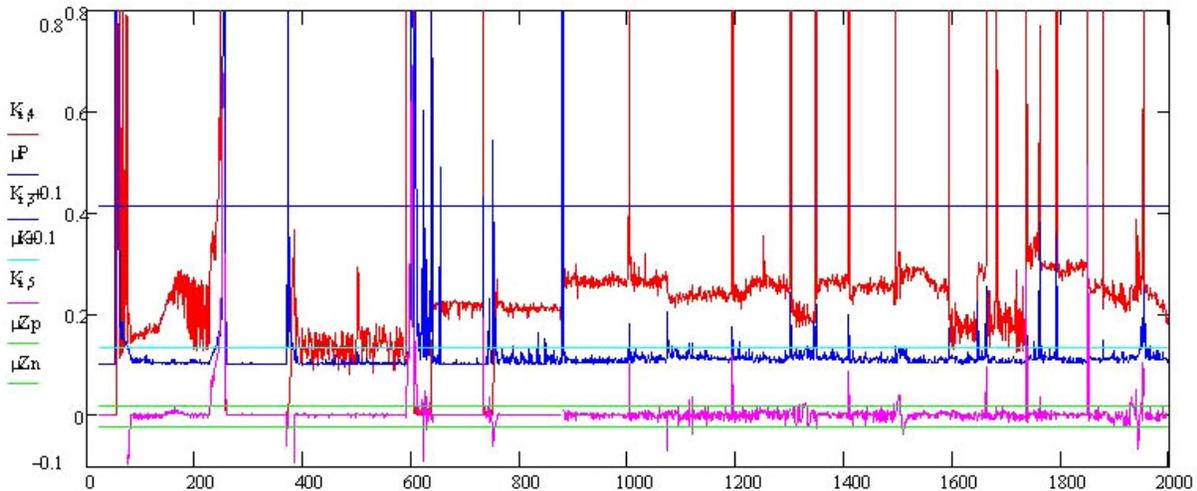
ключов кадър. При детекцията на прагови стойности може да се ползват и диференциални методи, каквито се ползват при анализа на физиологични сигнали (A10). Това е така защото в рамките на една визуална сцена е възможно да има движещи се обекти, което не означава че имаме начало на нова визуална сцена.

$$Mcl_i = \begin{cases} 1 & \text{за } (mclF_{i-1} \neq mclF_i) \wedge (M_i^* = 1) \\ 0 & \text{за } mclF_{i-1} = mclF_i \wedge (M_i^* = 1) \end{cases} \quad (17)$$

където: Mcl_i^* е индексен едномерен масив съдържащ единици само за онези ключови кадри, за които има промяна и на модата $mclF_i^*$ на цветната хистограма.

Използването на ключови кадри за анализ на видео филми и мултимедийни клипове се извършва на два етапа:

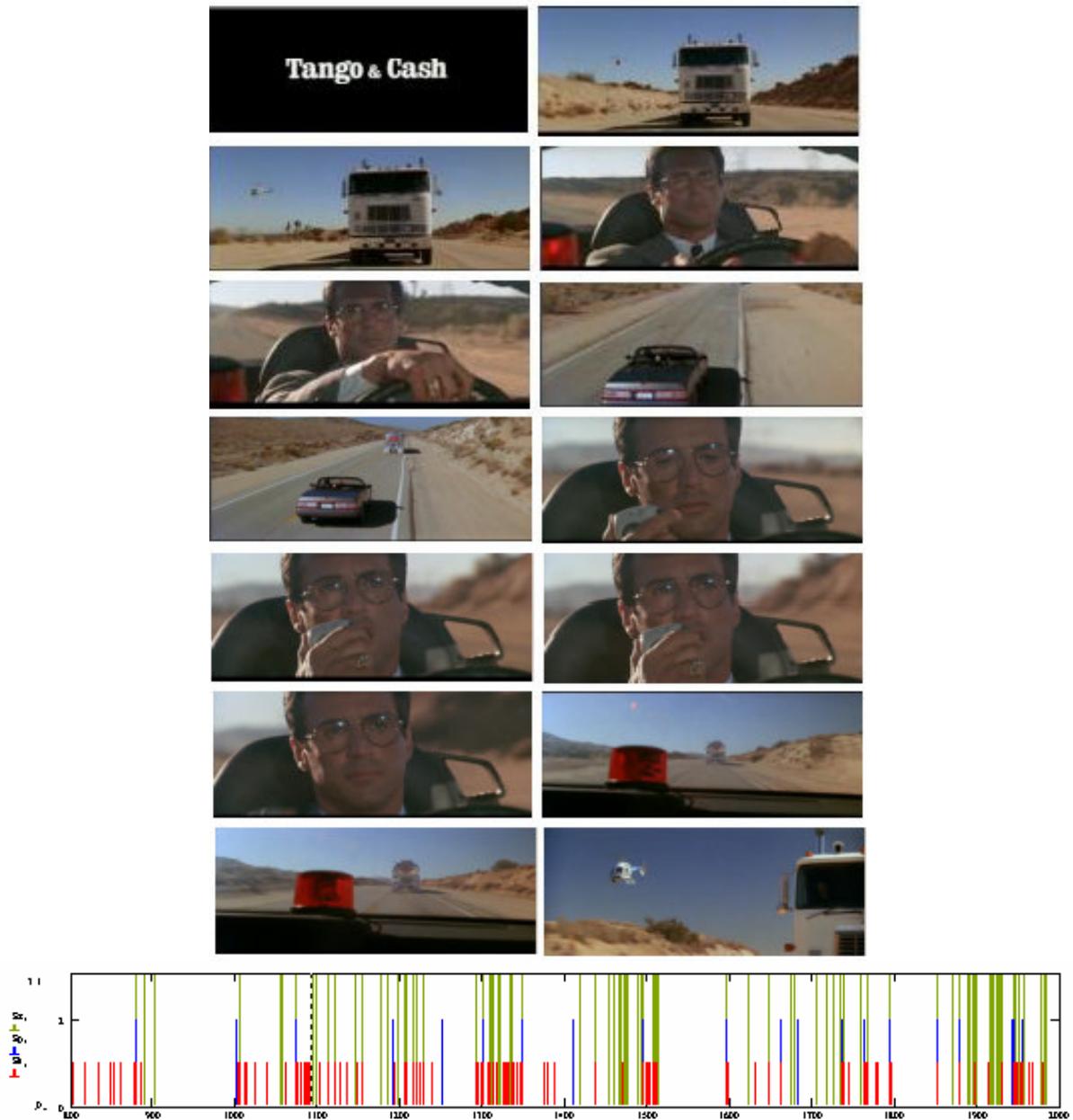
- **Детектиране на глобални ключови кадри** - използва се критерият на Пийрсън, като се търси отклонение над средното за цялата изследвана последователност. В тези области се търси нарастване на критерия на Колмогоров (фиг. 3.2);
- **Детектиране на локални ключови кадри посредством анализ на стойностите на критериите в рамките на една визуална сцена** – използват се комплексният критерий и критерият на Колмогоров.



Фиг. 3.2. Графика на статистическите критерии **K**-червените стълбчета, **X**-сините стълбчета и **Z**-зелени стълбчета за 2000 поредни кадъра;

Така детектираните ключови кадри се използват за автоматизираното сегментиране на последователностите от изображения в отделни семантично свързани визуални сцени. Индексацията се извършва чрез обхождане на индексните масиви с детектирани ключови кадри за трите вероятно статистически критерии (фиг. 3.2 и фиг. 3.6). Болшинството от грешките при детекция

на глобални ключови кадри са следствие от някакви локални дефекти в кадрите, напр. отблясъци наблюдавани само в един и същи кадър, предизвикани от наслагване на отделни дубли при монтажа на филмите. При индексирането на визуални сцени може да използваме и данни от аудио канала, анализиран чрез нелинейни методи (A10), например за нуждите на мониторинг на спортни предавания или новинарски емисии. Експерименталните резултати от сегментирането са осигурени чрез анализ на видео филми, като е извършено автоматизирано първоначално детектиране и броене на ключовите кадри и сцени посредством предложения алгоритъм (A8).



Фиг. 3.6 Сегментиране на 2000 поредни кадъра от филма “Танго и Кеш”

Така получените данни се сравняват с ръчно определените позиции на ключови кадри и визуални сцени (таблица 3.1).

Име на филма	Брой кадри в целия филм	Брой автоматично детектирани сцени
Navy	145	3
Tank IR	157	6
Rocket	159	4
Tango and Cash	146 919	2270
Rambo 4	130 900	1597
Train	158 979	1839

Таблица 3.1. Експериментални резултати от автоматизираното разделяне на видеофилми на отделни визуални сцени

Изводи към Глава 3

- Използването на вероятно статистическите критерии на Колмогоров, Пийрсън и комплексен критерий в анализа на последователностите от 3D хистограмите са гаранция за надеждно откриване на движение в реално време.
- При анализа на мултимедийна информация с цел по-точно семантично разделяне на визуалните сцени следва да се използват и данните от групирани цветни 2D хистограми.

Приноси към Глава 3

- Модифициран и подобрен е модел за детекция на движение ползващ критериите на Колмогоров, Пийрсън и Комплексен чрез анализ на 3D хистограмите.
- Създаден е модел за автоматизирано сегментиране на визуални сцени с детекция на локални и глобални ключови кадри отчитащ и модата на цветните хистограми.
- Експериментално е доказано, че критериите на Колмогоров и Пийрсън се ползват за детекция на глобални ключови кадри, а за Колмогоров и комплексен за локални ключови кадри.

Глава 4

МОДИФИЦИРАНЕ НА АЛГОРИТМИ ЗА 2D И 3D ХИСТОГРАМНА И ЕНТРОПИЙНА СЕГМЕНТАЦИЯ НА ПОЛУТОНОВИ ИЗОБРАЖЕНИЯ

Сегментиране на изображения чрез откриване на известни вероятностни разпределения в 3D хистограмата им

Тук е предложена модификация на 2D метода за сегментиране на изображения чрез откриване на подобни вероятностно статистически разпределения в техните 3D нормирани мощностни хистограми. При тези методи се постига висока прецизност при условие, че изображенията са заснети в константни условия, което е валидно за медицинските системи. Методите са по-бавни от ентропийната сегментация. Поради спецификата на подобен тип изображения и с цел осигуряването на възможност за сравняване на заснети с различна резолюция хистограми от сходни изображения, се използват нормирани мощностни хистограми (фиг. 4.2).

Изчисление на нормиращ коефициент за 3D хистограмата

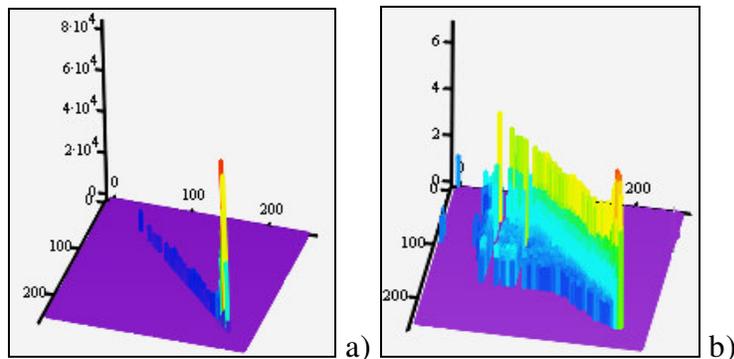
$$n = \frac{1}{\log(k * x * y)} \quad (18)$$

където: x и y са размерите на изображението

Преминаване към 3D мощностна хистограма

$$powerH3D_{i,j} = H_{i,j}^n \quad (19)$$

където: $powerH3D_{i,j}$ е мощностната 3D хистограма



Фиг. 4.2. 3D нормирана хистограма –(a), 3D нормирана мощностна хистограма – (b)

Разпознаването на обектите става чрез определяне на минималното разстояние отнесено към референтна 3D хистограма.

$$TARGET = \min \sum_{x=Xb*start}^{Xb*end} \sum_{y=Yb*start}^{Yb*end} m_{l(2D)} (H3D_{x,y} - p_{x,y})^2 \quad (20)$$

където: $H3D$ е примерната нормирана мощностна хистограма, $P_{x,y}$ е областта от сегмента на референтното разпределение за зоните (A, B, C и D) на ентропийно сегментираната 3D хистограма, $Xb*start, end$ и $Yb*start, end$ са съответно координатите на референтен набор, $m_{l(2D)}$ е математическото очакване за област 1

= A, B, C, D (виж Глава стр. 22)

Същият метод може да се използва за обработка на визуални сцени чрез анализ на техните 4D хистограми. Комбинацията на този подход с използването на обектните маски на най-значимите подвижни и стационарни обекти в сегментирани визуални сцени улесняват процеса на сегментиране, търсене и сравняване на визуални сцени.

Сегментиране на изображения чрез 2D ентропийна функция

Разгледани са и модифицирани методи за квантуване на изходните изображения на две и повече нива чрез многократна ентропийна сегментация на 2D и 3D хистограмите им (A11). Предложени са и реално имплементирани са функционални модели за хомогенизиране и обработка на изображения чрез групиране в 3D хистограмата и 3D ентропийните им функции. Предложени са нови комплексни процедури за многомерно детектиране и проследяване на подвижните обекти, тяхната класификация и отделянето на фона от целите. За определяне на 2D (Kapur J.N., P.K. Sahoo, A.K.C.Wong, 1985.) ентропийната функция работим с 2D хистограмата на дадено изображения.

$$P_{(2D)} = \{p_0, p_2, \dots, p_b, \dots, p_{255}\} \quad (21)$$

където: $P_{(2D)}$ са нормализираните хистограми на изображенията

p – е вероятността да се срещне пиксел със съответна яркост – b , а b приема стойности от 0 до 255.

Изчислението на ентропията става за всеки един стълб на $P_{(2D)}$,

като ентропията на изображението и се записва в $H(P_{(2D)})$:

$$H(A) = -\sum_{i=0}^b p_i \log p_i \quad (22)$$

$$H(B) = -\sum_{i=b}^{255} p_i \log p_i \quad (23)$$

$$H_b = -\log P(A) - \log P(B) - \frac{H(A)}{P(A)} - \frac{H(B)}{P(B)} \quad (24)$$

където: b = се мени от 0 до 255, A от 0 до b , B от 255 до b .

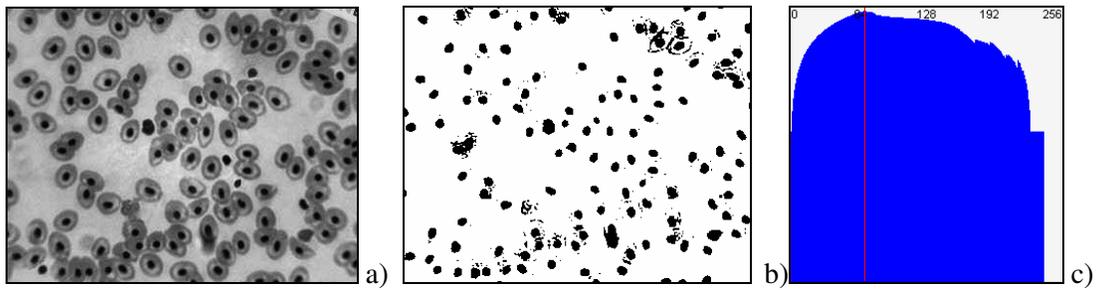
$$H(P_{(2D)}) = \{H_0, H_1, \dots, H_{255}\} \quad (25)$$

Определяне на максимума на ентропията:

$$b^* = \max H(P_{(2D)}) \quad (26)$$

Максимумът на ентропията определя стойността – b^* , която се използва като прагова за квантуването на изходното 8 битово

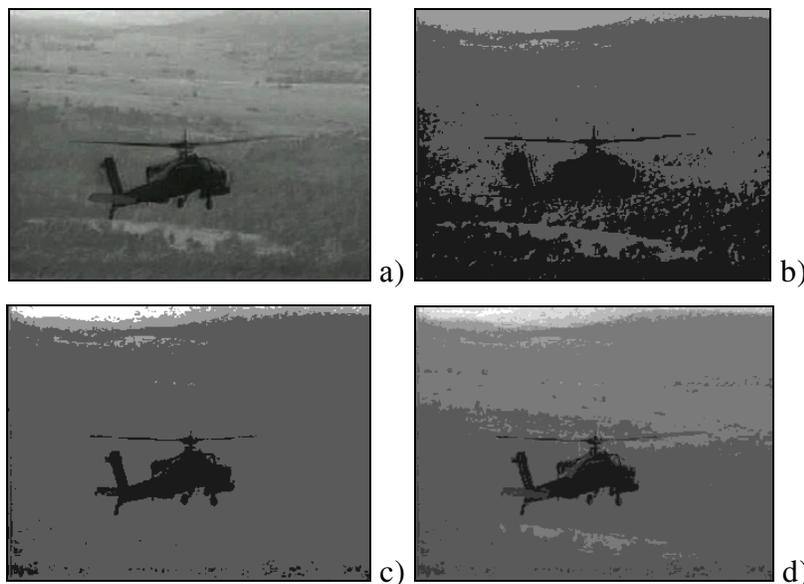
изображение на 2 нива – черно и бяло (фиг. 9 - пример със сегментация на изображение с биологични клетки). В подобни случаи квантуването на 2 нива не дава добри резултати, особено без филтрация за отстраняване на фона на изображенията.



Фиг. 9. Биологични клетки – a), квантуване на две нива с оптимален праг на квантуване –b) чрез изчисление максимума на 2D ентропийната функция -c)

Адаптивно ентропийно сегментиране на 4 и 8 канала

Разработено е адаптивно ентропийно квантуване на 4 и 8 канала, чрез многократно откриване максимумите на 2D и 3D ентропийната функция на изображенията и нейното делене (фиг. 4.8).



Фиг. 4.8 Изходно изображение - a), квантувано на 4 нива с фиксирани прагове - b), квантувано на 4 нива с локални ентропийни максимуми - c) и квантувано на 8 нива - d) с локални ентропийни максимуми.

- Изчисляваме b^* от $H(P_{(2D)})$, $b = \text{var from } 0 \text{ to } 255$
- Създаваме две еднакви хистограми $P1_{(2D)}$ и $P2_{(2D)}$, изпълнявайки долното правило:

$$p1_b = \begin{cases} p_b & \text{if } b^* \leq \max \\ 0 & \text{if } b^* > \max \end{cases} \quad (27)$$

$$p2_b = \begin{cases} p_b & \text{if } b^* \geq \max \\ 0 & \text{if } b^* < \max \end{cases} \quad (28)$$

където $b = \text{var}$ от 0 до 255

- Изчисляване на $H(P1_{(2D)})$ и $H(P2_{(2D)})$
- Изчисляване на b_1^* за $P1_{(2D)}$ и b_2^* за хистограма $P2_{(2D)}$
- b_1^* , b^* , b_2^* определят позициите на максимума на ентропийната функция и двете локални ентропийни функции на изображението, изчислени за $P1_{(2D)}$ и $P2_{(2D)}$, тези стойности ползваме за квантуване изображението на 4 нива (Фиг. 10).
- Можем да продължим сегментацията на така обособените области в 2D хистограмата, като квантуваме изображението на 8 нива.

Сегментиране на изображения чрез максимумите в 3D ентропийната функция

Разгледани са два метода за сегментация на изображенията чрез анализа на техните 3D ентропийни функции (A9). След което се извършва повторна сегментация за така разделените зони (фиг. 4.9).

Определяне на пълната вероятност за областите A и B :

$$P_{rr} = \sum_{i=0}^r \sum_{j=0}^r p_{ij}, \quad (29)$$

$$1 - P_{rr} = \sum_{i=r+1}^k \sum_{j=r+1}^k p_{ij} \quad (30)$$

Определяне на ентропията за областите A , B и цялата хистограма:

$$E_{rr} = \sum_{i=0}^r \sum_{j=0}^r p_{i,j} \ln p_{i,j} \quad (31)$$

$$E_{r+1,r+1} = \sum_{i=0}^r \sum_{j=0}^r p_{i,j} \ln p_{i,j} \quad (32)$$

$$E_k = \sum_{i=0}^r \sum_{j=0}^r p_{i,j} \ln p_{i,j} \quad (33)$$

Условно разделяне на вероятностното разпределение на две части A и B с P_{pp} и $1 - P_{pp}$, съответно:

$$P(A) = \frac{P_{00}}{P_{pp}}, \frac{P_{01}}{P_{pp}}, \dots, \frac{P_{pp}}{P_{pp}} \quad (34)$$

$$P(B) = \frac{P_{p+1,p+1}}{1 - P_{p+1,p+1}}, \frac{P_{p+1,p+2}}{1 - P_{p+1,p+1}}, \dots, \frac{P_{p+k,p+k}}{1 - P_{p+1,p+1}} \quad (35)$$

Ентропите за частите A и B на хистограмите на изображението се изчисляват както следва:

$$H(A) = \sum_{i=0}^r \sum_{j=0}^r \frac{P_{ij}}{P_{rr}} \ln \frac{P_{ij}}{P_{rr}} = -\frac{1}{P_{rr}} \left[\sum_{i=0}^r \sum_{j=0}^r (P_{i,j} \ln P_{i,j} - P_{r,r} \ln P_{rr}) \right] = \ln P_{r,r} + \frac{H_s}{P_{rr}}$$

$$H(B) = \sum_{i=r+1}^k \sum_{j=r+1}^k \frac{P_{ij}}{1 - P_{rr}} \ln \frac{P_{ij}}{1 - P_{rr}} = -\frac{1}{1 - P_{rr}} \left\{ \sum_{i=r+1}^k \sum_{j=r+1}^k [p_{i,j} \ln p_{i,j} - (1 - P_{r,r}) \ln(1 - P_{rr})] \right\} =$$

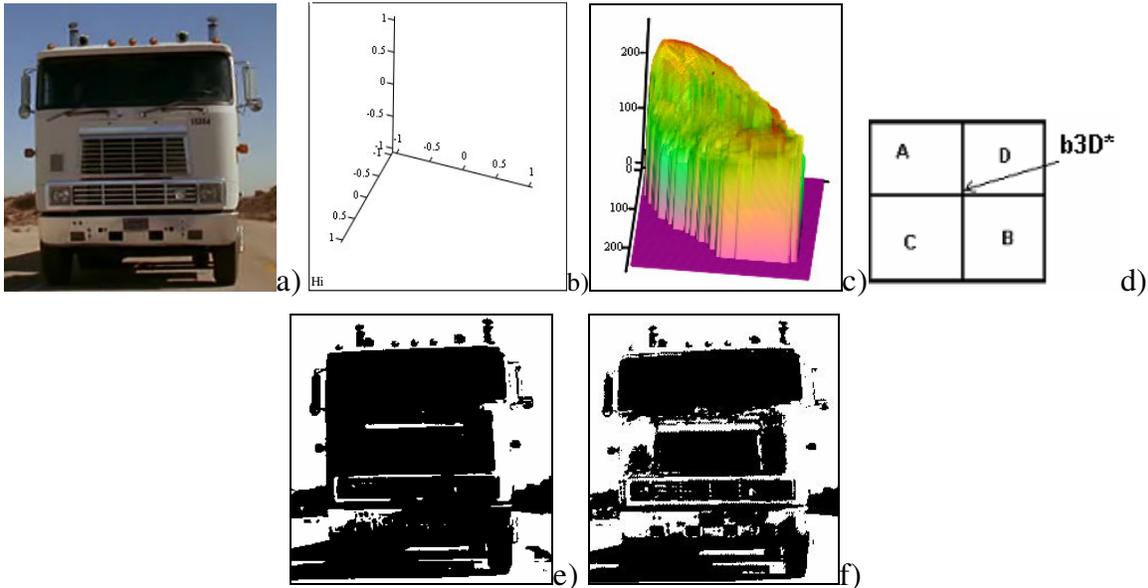
$$= \ln(1 - P_{r,r}) + \frac{H_s}{(1 - P_{rr})} \quad (36)$$

Сумарният израз за частите A и B определя функцията $f(rr)$

$$f(rr) = H(A) + H(B) = \ln P_{rr} (1 - P_{rr}) + \frac{H_{rr}}{P_{rr}} + \frac{H_{kk} - H_{rr}}{1 - P_{rr}} \quad (37)$$

Максимумът на ентропийната функция дефинира оптималния праг на квантуване:

$$b^* = \max f(r, r) \quad (38)$$

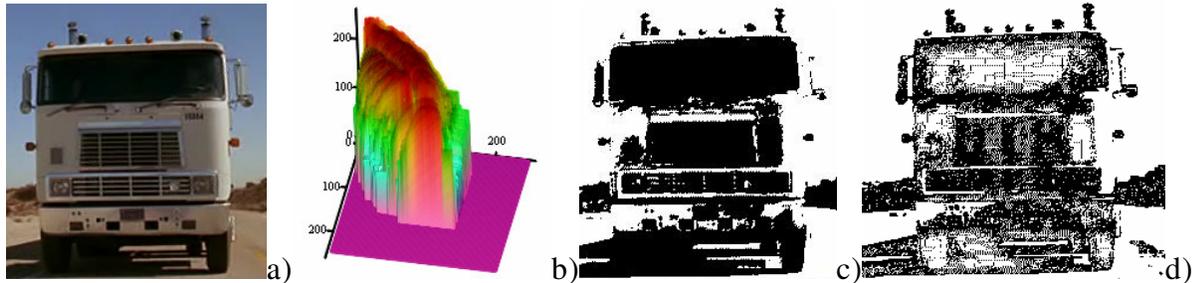


Фиг. 4.9. Оригинално изображение – а), 3D хистограма – б), 3D ентропия – с), която дели 3D хистограмата на 4 зони А, В, С, D – д), изображение квантувано с 2D ентропийна функция – е), 3D ентропийно изображение – ф)

3D ентропия в интервалите на разликата между яркостта на всеки хистограмен стълб спрямо всички останали яркости

Предложен е метод за сегментация с 3D ентропия между

яркостите на две съседни области (A11), който може да се използва за изображения, имащи равностойни по форма и яркост обекти и такива, които са предварително сегментирани с 2D ентропийна функция. При прилагане на функцията в последователност от изображения става възможно откриването на най-бързото нарастване на функцията в последователност от 3D хистограми. Тази ентропия наричам – градиентна. 3D градиентната ентропия се изчислява за всички отделни канали (ентропийни слоя) на изображението (фиг. 4.10).



Фиг. 4.10 Оригинално изображение – a), 3D градиентна ентропия – b), 3D ентропийно изображение – c), 3D градиентно ентропийно изображение – d).

Този тип изчисление може да отчита диагонала на хомогенните зони в 3D хистограмата или да работи без него. В този случай се постига сегментиране само на зони с яркостен преход, като хомогенните зони се отстраняват. Приложение на метода има при сегментацията на движение на визуални сцени с 4D ентропийна функция, описано в Глава 5.

Векторен модел за класификация на обектите

След многоканална ентропийна сегментация на визуалните сцени може да пристъпим към много параметрично обектно сегментиране, като определяме следните параметри на всеки един обект (A1).

Оптико-геометричните параметри на обектите:

- хоризонтален и вертикален размер Z_x и Z_y ;
- площ S ;
- периметър P ;
- коефициент на формата k_ϕ ;
- център на тежестта;

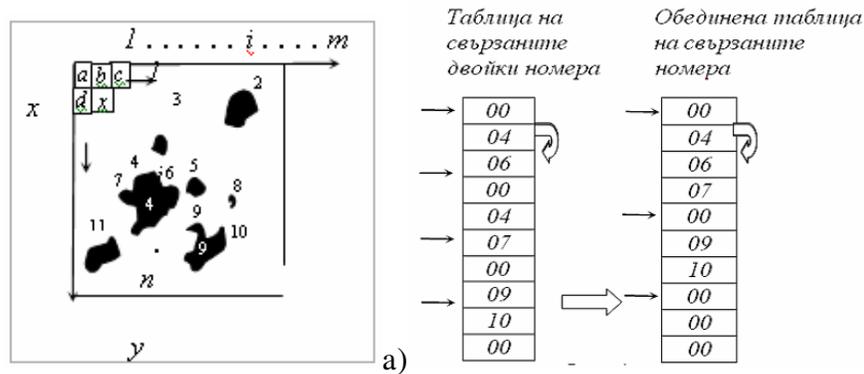
Допълнителни параметри на обектите:

- оптическа плътност и коефициент на отклонение;
- 2D и 3D локални хистограми на обектите;
- работа с цветните хистограми;

Процедура за маркиране и броене на обектите

Предложена е процедура за маркиране и борене на обектите в

ентропийните маски на изображенията чрез двукратно сканиране на цялото изображение (фиг. 4.13) и създаването на индексирани свързани таблици. Алгоритъмът се отличава от масово прилаганите разновидности на запълващи процедури.



Фиг. 4.13. Оператор за броенето – а), и създаването на свързани таблици, които се преиндексират по време на втората итерация на оператора.

Използване на цветния определител на изображенията

При работа с цветни изображения се изчислява тяхната сегментирана цветна хистограма (Bruce J. Tucker B. Manuela V. "Fast and Inexpensive Color Image Segmentation for Interactive Robots", 2005), като се ползва HIS (Hue Intensity Saturation) цветен модел. За целта HIS пространството се дели на сегменти от по 36^0 , групиращи всички цветове към следните основни: черен, сив, червен, кафяв, жълт, зелен, син, синьозелен, пурпурен и бял. Изчислява се цветен определител за цялото изображение:

$$CI_C = \frac{N_C}{S} \quad (39)$$

където: CI_C е цветен определител за отделния обект, или ентропийна маски или цялото изображение, N_C - брой пиксели с определен цвят, S - брой пиксели

Сравняването на CI_C за отделните изображения подобрява разделимостта, когато се анализират изображения с близки 3D хистограми и ентропийни маски. **Модата** на CI_C хистограмата се използва, като условие позволяващо слепване на две съседни визуални сцени, за които имаме детектиран ключов кадър съгласно праговите стойности на критериите на Пийрсън и Колмогоров. Подробностите касаят редукция на грешката при индексирание на визуални (Глава 6).

Векторен модел за представяне на визуални сцени

Всички разгледани процедури за сортиране и групиране на обектите изискват интензивни изчисления, целящи откриване на минимални разстояния между хистограми с параметрите на обектите.

За нуждите на този дисертационен труд са избрани най-често ползваните в практиката метрики:

- **Разстояние на Бхатачария**

$$B(V_Q, V_C) = -\ln \sum \sqrt{V_Q(i), V_C(i)} \quad (40)$$

където: V_Q и V_C са векторите които сравняваме, B е резултатното разстояние

- **Дивергентен фактор**

$$D(V_Q, V_C) = \sum_i \left[(V_Q(i) - V_C(i) \ln \frac{V_Q(i)}{V_C(i)}) \right] \quad (41)$$

където: D е резултатното разстояние

- **Евклидово разстояние**

$$L_2(V_Q, V_C) = \sqrt{\sum_i (V_Q(i) - V_C(i))^2} \quad (42)$$

където: L е резултатното разстояние

- **Разстояние на Матсушита**

$$M(V_Q, V_C) = \sqrt{\sum_i (\sqrt{V_Q(i)} - \sqrt{V_C(i)})^2} \quad (43)$$

където: M е резултатното разстояние

- **Метод на Сейн и Балард**

$$V(V_Q, V_C) = \frac{\sum_i \min(V_Q(i), V_C(i))}{\min(\sum_i V_Q(i), \sum_i V_C(i))} \quad (44)$$

където: V е резултатното разстояние

Векторно пространство на визуална сцена

Формирането на векторното пространство се извършва в следната последователност:

- Предварителна обработка – филтрация и хомогенизиране.
- Ентропийна сегментация на 3D хистограмите за всяко едно изображение, след което отделните ентропийни маски преминават през процедура за маркиране и броене на обектите;
- Определяне на оптико-геометричните параметри на обектите и цветния определител на изображенията. На този етап можем да извършим параметрична филтрация на отделните изображения;

- Изчисляване на обобщена 2D обектна хистограма, съдържаща данни за броя обекти по канали, максимални по размер и коефициент на формата обекти, средно значение и стандартно отклонение от размера, формата, максимален цветен определител и девиация на центровете на тежестта на обектите за всеки един ентропиен канал;
- На база референтния фонен кадър можем да извършим сегментация на движението чрез анализ на 3D хистограмата на всяко едно изображение с хистограмата на фонения кадър;
- Създаваме многомерен вектор за всяка една визуална сцена, съдържащ ентропийно сегментираната 4D хистограма, 2D хистограмата на отделните определители за всеки един ентропиен канал, векторното пространство на движението, данни за метриката на качеството;

Изводи към Глава 4

- Вероятностните методи за разпознаване на цифрови изображения намират приложение в областта на медицината и където условията на заснемане са константни.
- Използването на 3D ентропийната функция за сегментация на изображения позволява по-оптималното групиране на яркостните преходни зони в изображението.
- Използването на 3D градиентната ентропия позволява едновременното групиране на яркостни преходи и отстраняване на фона на изображенията.
- При работа с цветни изображения се постига по-добро семантично разделяне на визуалните сцени на база мода на цвета.

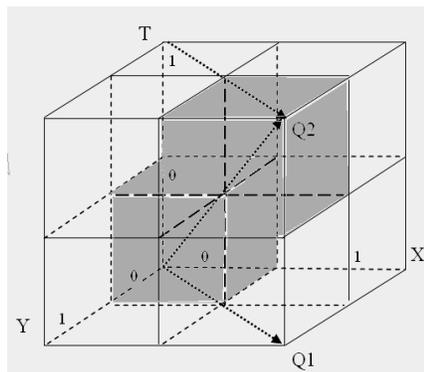
Приноси към Глава 4

- Разработен е метод за сегментация на изображения чрез многократно ентропийно групиране на 2D и 3D хистограмите им и 3D градиентна ентропия.
- Реализиран е много параметричен векторен модел за представяне на отделните изображения.

Глава 5

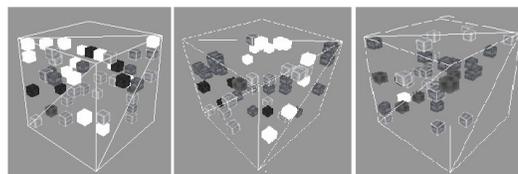
РАЗРАБОТКА НА ФУНКЦИОНАЛЕН МОДЕЛ ЗА БЛОКОВА 4D ХИСТОГРАМНА ОБРАБОТКА НА СЕМАТИЧНО СВЪРЗАНИ ПОСЛЕДОВАТЕЛНОСТИ ОТ ИЗОБРАЖЕНИЯ

Предложеният тук 4D хистограмен метод позволява блоковата вероятно статистическата обработка на визуални сцени. При обработката на последователностите от изображения с използване на 4D ентропийна функция се постига бърза сегментация на движението (A5). Едновременно с това процесът изисква наличието на по-сериозни хардуерни ресурси отколкото 3D хистограмните методи. Предимство на предлагания подход е, че се обезпечава едновременно обработката на информацията от цялата визуална сцена. Чрез това се постига бърза хомогенизация и сегментация в цялата последователност. 4D хистограмата има диагонал на повтарящите се във времето хомогенни зони – $(0, Q2)$, равнина на яркостните разпределения в пространството на 2D изображенията – $(0, T, Q2, Q1)$, като всички останали области може да бъдат третирани като шум (фиг. 5.2).



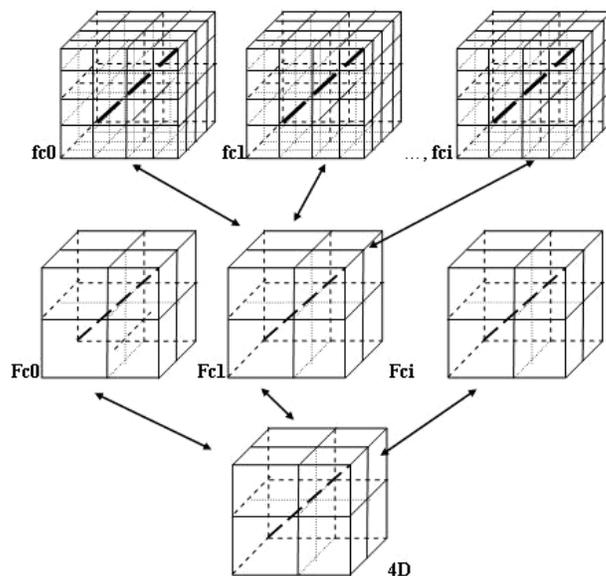
Фиг. 5.2 1 битова 4D хистограма (8 кубчета с различна плътност)

Анализът и съхранението на 8 битови 4D хистограми може да е и бавно, като отнема прекалено много памет (над 16MB). Затова предложеният 4D хистограмен модел за обработка на визуални сцени използва предварително ентропийни сегментирани 2 и 3 битови изображения на 4 и 8 нива.



Фиг. 5.4. Дадени са 3 независими 4D хистограми, като целта е да бъдат открити други визуални сцени с подобни 4D хистограми.

Това ограничава обема памет, заеман от 4D хистограмата до 512 байта. Чрез 4D ентропийна сегментация е възможно многокритериалното им индексване (фиг. 5.4). При преминаването към 4D мощностни нормирани хистограми е възможно сравняването на визуални сцени с различна продължителност и резолюция (фиг. 5.5).



Фиг. 5.5. Двуслойна архитектура за класифициране на визуални сцени с 4D хистограми. $Fc1, Fc2, \dots, Fci$ – са нормирани индексирани 1 битови 4D хистограми, $fc1, fc2, \dots, fci$ – са нормирани 2 битови 4D хистограми.

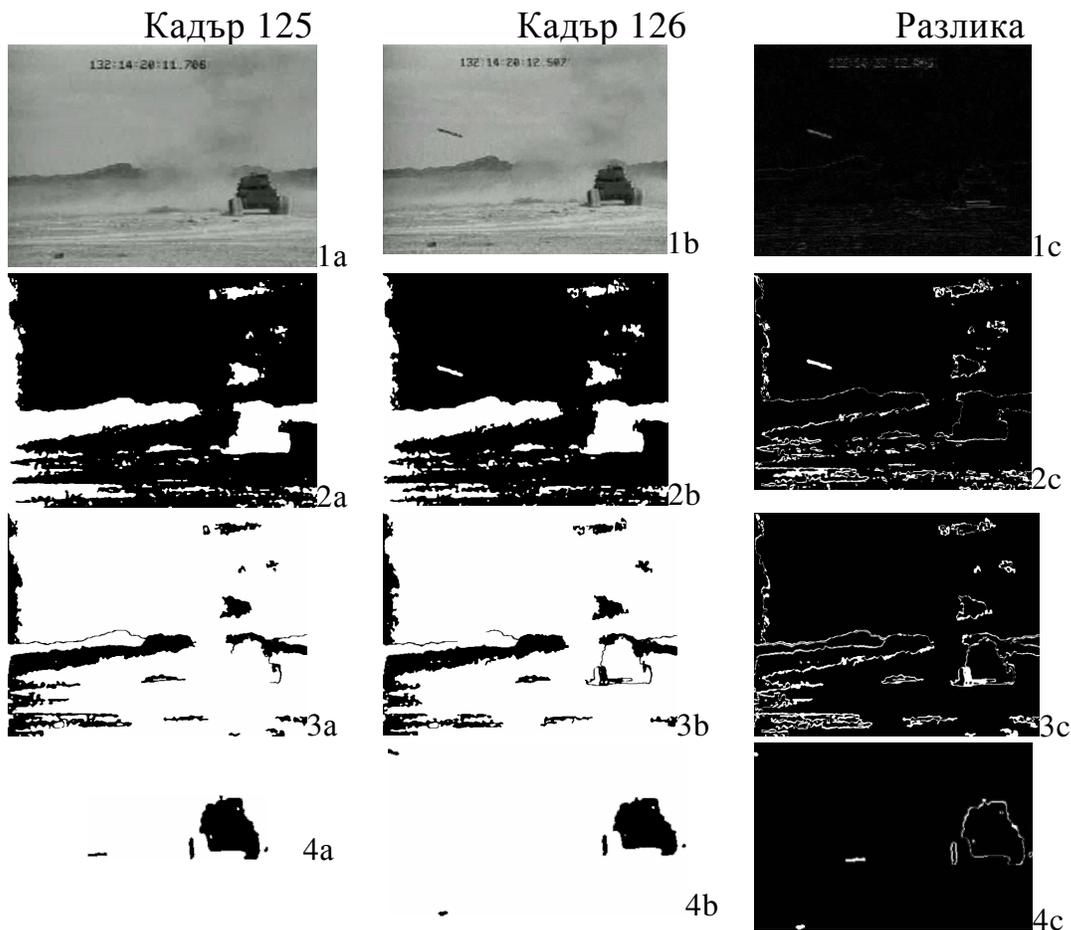
Хистограмен метод за многоканална ентропийна сегментация на последователности от изображения

Методът използва процедури за броене на обектите в отделните ентропийни канали на последователността от изображения (фиг. 17). Отчитането на измененията става посредством следене на броя на обектите в отделните ентропийни маски и сравняване с предходните кадри (A1). Комбинацията на 4D хистограмния модел може да се ползва като допълнителен параметър характеризиращ визуалната сцена (брой канали, брой обекти в каналите, среден размер на обектите в каналите и други) (таблица 5.1).

Данни за кадър 125	Данни за кадър 126
Канал 1 – 1 обект (фиг. 17.2a)	Канал 1 – 1 обект
Канал 2 – 16 обекта (фиг. 17.3a)	Канал 2 – 11 обекта
Канал 3 – 3 обекта (фиг. 17.4a)	Канал 3 – 6 обекта

Таблица. 5.1 Брой обекти в трите канала на изображенията от Фиг. 17.

Броят на обектите в ентропийно сегментираните последователности от изображения позволява дефинирането на допълнителен критерий за разделимост на визуални сцени, когато вероятно статистическите критерии и модата на цветната хистограма дават голям брой грешки, или когато имаме движещи се обекти с близки вероятно статистически параметри, като цвят и яркост.

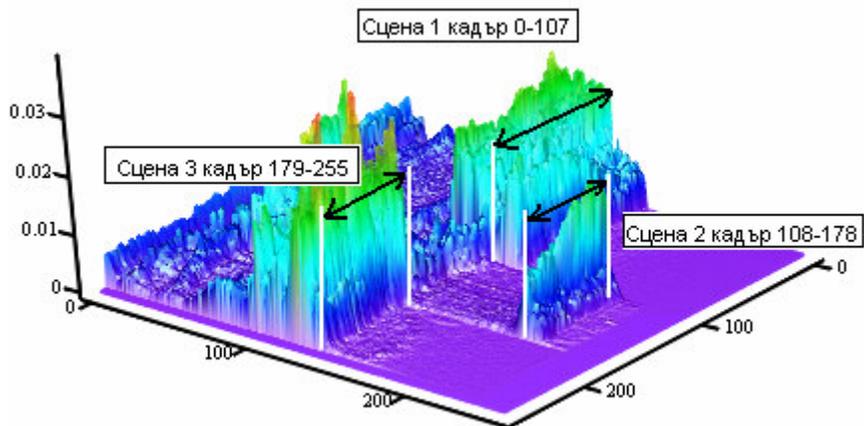


Фиг. 5.6. Многоканален ентропийен анализ на последователности от изображения, чрез броене и изчисление площта на обектите в отделните маски

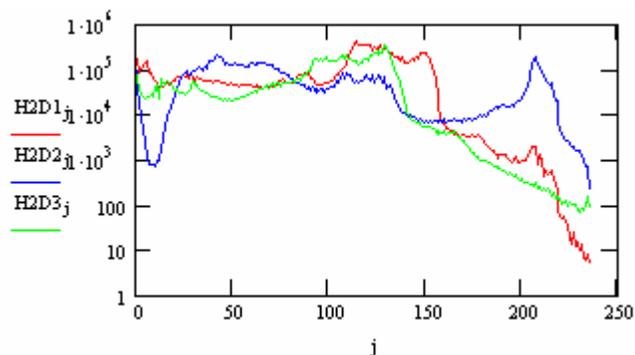
При формиране на масива от 2D хистограми за отделните ентропийни маски от всички изображения в дадена визуална сцена е възможно изчисление на осреднената и 2D хистограма (фиг. 5.7.b). Разбира се това може да се приложи и за изчисление на референтна 3D хистограма за всяка една визуална сцена. Тъй като обработката на видео филми

$$H2D_j = \sum_{t=start}^{end} H2D_{t,j} \quad (45)$$

където: H2D е глобалната мощностна 2D нормирана хистограма на последователността в интервала start – end, t е номера на всеки пореден кадър, j = 0 до 255 (0 до 8 за ентропийно сегментирано изображение) отговаря на поредния стълб в 2D хистограмата (Фиг. 18.a)



a)



b)

Фиг. 5.7 Графично представяне на 255 поредни 2D нормирани хистограми от 3 последователни визуални сцени –a), осреднени 2D хистограми за трите визуални сцени ($H2D1$, $H2D2$ и $H2D3$), нормирани в логаритмичен мащаб – b).

Разпределение на обектите по канали за цялата последователност от изображения

Ако предположим че един филм може да се разложи на N сцени, за всяка една такава последователност се формират $N1$ вътрешни ключови кадри.

$$OF = \sum_{i=0}^K O_i, \quad O_i = \begin{cases} 1 & \text{for } O_i = \text{target} \\ 0 & \text{for } O_i = \text{noise} \end{cases} \quad (46)$$

където: K е общия брой обекти, OF е броят обекти в 1 ключово изображение, target е условието, на което трябва да отговаря много параметричният вектор на обектът O_i , за да бъде отчетен във всяко едно изображение

$$OS = \sum_{i=0}^M OF_i, \quad OF_i = \begin{cases} 1 & \text{for } OF_i = \text{target} \\ 0 & \text{for } OF_i = \text{noise} \end{cases} \quad (47)$$

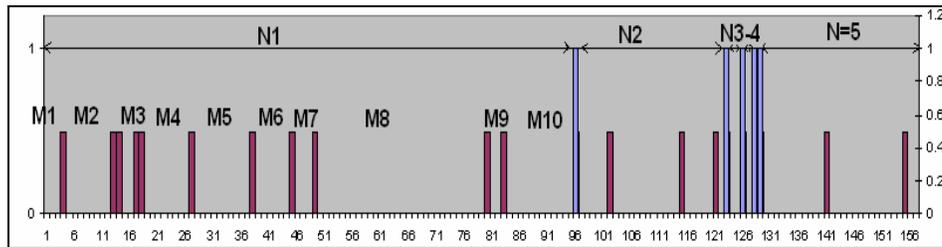
където: OS е броят на обектите в 1 ключов интервал M , target е условието, на което следва да отговарят обектите в рамките на една визуална сцена

Условията target и noise се изчисляват адаптивно след сортиране

и параметрична филтрация на обектните вектори, като е възможно и ръчно задаване на условия за филтрация (площ, форма, яркост, цвят).

$$ON = \sum_{i=0}^N OS_i \quad (48)$$

където: **ON** – броят обекти във всичките **N** сцени (фиг. 5.8).



Фиг. 5.8. Разбиване на целия филм на отделни сцени посредством значимите ключови кадри – между 2 пикови кадъра се образува една сцена, и намирането на ключовите кадри в отделните сцени посредством пиковите зони.

Изводи към Глава 5

- Обработката на последователности от семантично свързани изображения може да се осъществи блоково чрез 4D хистограма.
- Комбинирането на данните от 4D и 2D многоканалния обектен анализ позволява създаването на система за бърза индексация, търсене и съпоставяне на визуални сцени с различна продължителност и заснети с различна резолюция.
- Време пространствената 4D ентропийната функция може да се използва при формирането на референтен фонев кадър.

Приноси към Глава 5

- Разработен е комбиниран 4D хистограмен модел за обработка и анализ на последователности от изображения, включващ и данните от обектната сегментация.

Глава 6

МЕТОДИКА ЗА ОТКРИВАНЕ НА ГРЕШКАТА ПРИ АВТОМАТИЗИРАНО ДЕТЕКТИРАНЕ НА КЛЮЧОВИ КАДРИ И МЕТРИКА НА КАЧЕСТВОТО НА ЦИФРОВИТЕ ИЗОБРАЖЕНИЯ

Методика за определяне на грешка при сегментация на визуални сцени

Детектирането на ключови кадри при рязка смяна на информационното съдържание в две последователни изображения

еднозначно указва за начало на нова визуална сцена. Не така обаче стои проблемът със сегментиране на визуални сцени с бавно изменящ се фон или бавно движещ се обект. В тези случаи се налага да бъдат използвани и допълнителни данни като глобалната цветна и обектна хистограми в ентропийно сегментираните изображения. При анализа на точността за детекция на ключови кадри са предложени индексите:

- **Брой автоматично детектирани ключови - ТКФС кадри, съвпадащи с ръчно потвърдените ключови кадри;**

$$TKFC = \sum_{i=Fstart}^{Fend} \begin{cases} 1 & \text{if } AK_i = HK_i \\ 0 & \text{if } AK_i \neq HK_i \end{cases} \quad (49)$$

където: **TKFC** (True Key Frames Count) е броят истинските ключови кадри в интервала $Fstart$ до $Fend$, **AK_i** е индексен масив с позициите на автоматично сегментираните ключови кадри, **HK_i** е индексен масив на ръчно определените позиции на ключови кадри

- **Индекс на изпуснатите ключови кадри - МКФИ;**

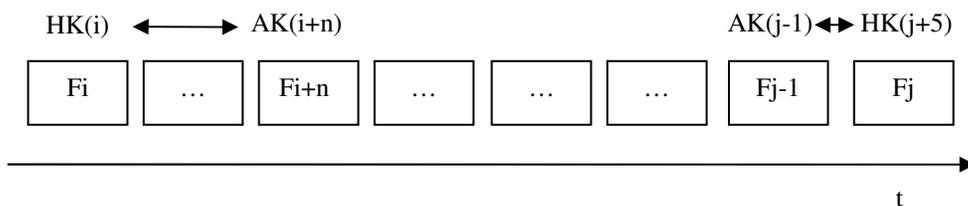
$$MKFI = \frac{\sum_{i=Fstart}^{Fend} AK_i}{\sum_{i=Fstart}^{Fend} HK_i} \quad (50)$$

където: **MKFI** (Missed Key Frame Index) указва отношението между общия брой автоматизирано открити ключови кадри към общия брой ръчно потвърдени ключови кадри.

- **Средно кадрово отместване**

Този тип грешка позволява да определим доколко автоматизираната система успява да установи движението преди оператора (фиг. 6.1). Методиката предполага изчисляване на средните значения за тези индекси, които нарастват при следните обстоятелства (таблица 6.1).

- прекалено бързо движение на обекта и фона - обикновено се наблюдава при анализ на видео филми и спортни предавания;
- прекалено бавно предвижване на обекта или когато обектът не е достатъчно представителен при системите за охрана и наблюдение.



Фиг. 6.1. Автоматично детектирани ключови, HK_i – е позицията на ръчно определен ключов кадър, AK_i – е позицията на автоматично открит, F_i , е пореден кадър с индекс i , $j-1$.

$$lkfd_i = \min\{(i+n) - i; (i+n) - j\} \quad (51)$$

където: **lkfdi** – (Local Key Frame Distance Index) е броят кадри показващ отместването на автоматично детектиран ключов кадър от всеки един ръчно индексирани ключов кадър.

$$KFDI = \begin{cases} 0 & \text{if } MKFI = 1 \\ \frac{1}{MK} \sum_{i=0}^N lkfd_i & \text{else if} \end{cases} \quad (52)$$

където: **KFDI** (Key Frame Distance Index) е средното отместване на детектираните ключови кадри от ръчно индексирани, **MK** (Missed Key) пропуснати ключови кадри, **i** е поредният кадър, **lkfd_i** е отместването

Име на филма	Брой кадри	TKFC	MKFI	KFDI
Navy	145	3	0.6	12.50
Tank IR	157	5	1.2	6.00
Rocket	159	4	1.0	0.00
Tango and Cash	146 919	2120	1.0708	0.04
Rambo 4	130 900	1430	1.1168	0.03
Train	158 979	1750	1.0509	0.07

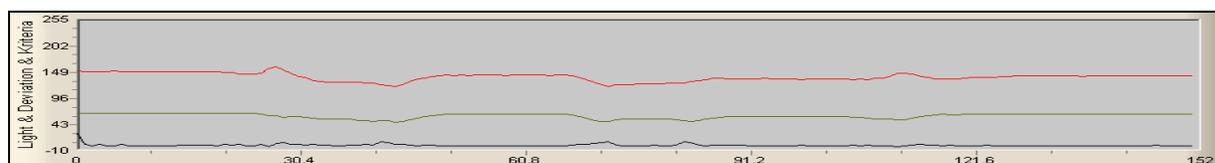
Таблица 6.1. Изчисление индексите на детектирани ключови кадри, с използване на корекция на база модата на цветната хистограма.

Детекция на движение в ентропийно сегментирани последователности от изображения

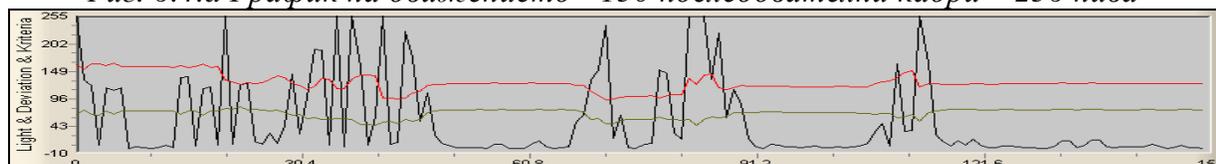
Сегментирането на движението се извършва и чрез анализа на ентропийно квантувани изображения на 2, 4, 8, 16 и 32 нива. При това детектирането на ключови кадри може да има известно разминаване от 2-3 кадъра в сравнение с детектираните ключови кадри при 8 битово изображение (фиг. 6.3.а) (A9). Алгоритъмът добре реагира на малки и големи изменения при 2 и 3 битови изображения (фиг. 6.4.а и 6.4.с).



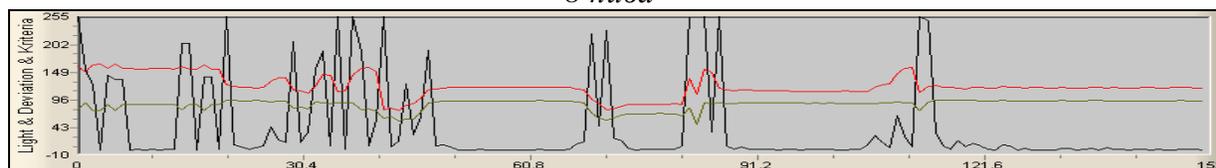
Фиг. 6.3 Оригинални изображения – горе, и черно бели изображения получени чрез квантуване на 2 и 4 и 8 нива, с оптимален праг на квантуване.



Фиг. 6.4.a График на движението - 150 последователни кадри – 256 нива



Фиг. 6.4.b График на движението на 150 последователни кадри - сегментация на 8 нива



Фиг. 6.4.c График на движението на 150 последователни кадри - сегментация на 4 нива



Фиг. 6.4.d График на движението 150 последователни кадри - сегментация на 2 нива

Методи за оценка на качеството на цифрови изображения

При анализ на качеството на цифрови изображения се използват модели, отчитащи особеностите на човешкото зрение. Подобни модели са отчетени в следните рекомендации: ITU-R Rec. BT.500-11 (2002), дележащи качеството на изображенията на следните категории: *отлично, добро, задоволително, слабо и лошо*.

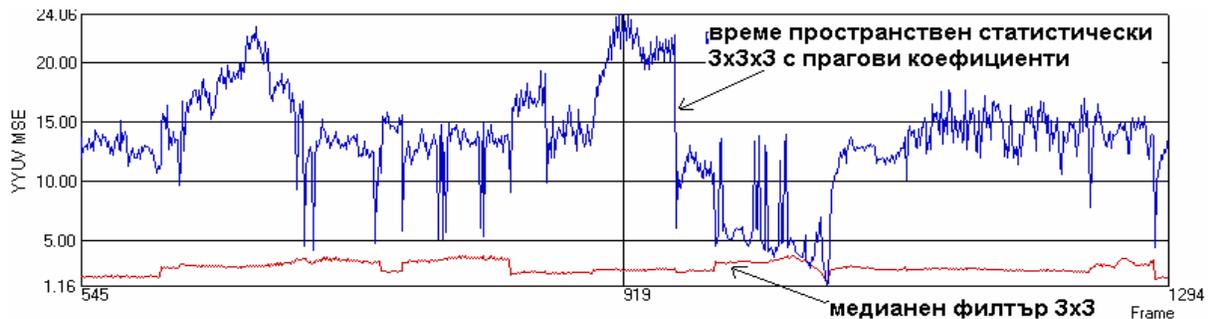
Метрика на качеството

Тези изчисления се правят преди и след обработка на изображенията и сегментация на визуалните сцени, обектна филтрация и ентропийна сегментния.

Средноквадратичната грешка е най-често използваният метод за определяне на грешката при обратна трансформация на компресирани изображения или след цифрова или обектна филтрация (фиг. 6.5).

$$MSE = \frac{1}{TXY} \sum_t \sum_x \sum_y \left[I(y, x, y) - \tilde{I}(t, x, y) \right]^2 \quad (53)$$

където: MSE – средноквадратична грешка, X и Y са размерите на изображенията, T брой кадри



Фиг. 6.5 Изчисление на MSE - за филтрирани с време-пространствен статистически и пространствен медианен филтър.

$$RMSE = \sqrt{MSE} \quad (54)$$

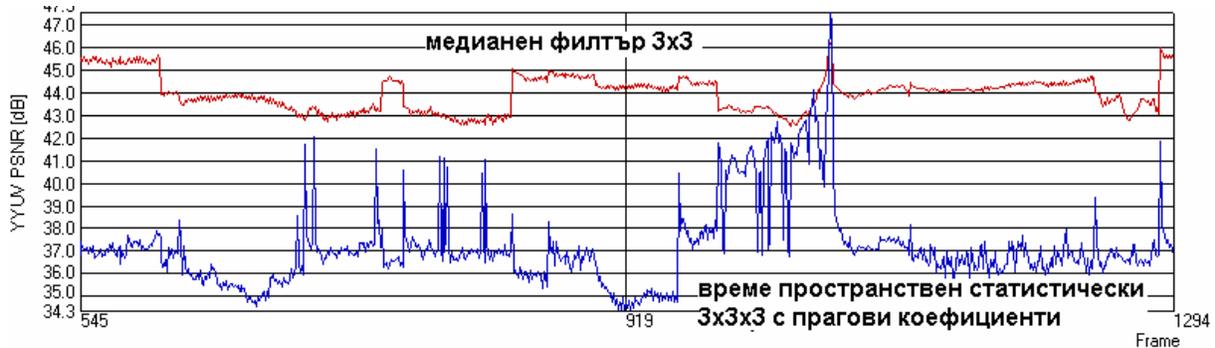
където: PSNR е пиковото отношение сигнал/шум

$$PSNR = 10 \log \frac{m^2}{MSE} \quad (55)$$

където: PSNR се мери в dB, m е 2^n за n битово изображение.

Метриците **PSNR** (Фиг. 6.6) и **MSE** (Фиг. 6.5) могат да се използват само за анализ на сиви полутонови изображения или поотделно в съответните RGB канали. По-обективно е прилагането на метриката за HSI цветен модел, тъй като по-добре отразява особеностите на човешкото зрение. Освен тази метрика се препоръчва

ползването на критерия на **Пийрсън**, който добре реагира при детекция на съществени разлики в двете изображения, което свидетелства за силно понижено качество.



Фиг. 6.6 Изчисление на PSNR за филтрирани с медианен и време пространствен статистически филтри. За анализа на качеството на обработените видео последователности е използван MSU Video Quality Measurement Tool

Използване на контраста за определяне на качеството на изображенията

Човешкото око е силно чувствително към изменения в осветеността, както и към локалните изменения на яркостите. Това обосновава използването на контраста при анализ на качеството на черно белите изображения. Контрастът се дефинира чрез:

$$C = \frac{Y_{\max} - Y_{\min}}{Y_{\max} + Y_{\min}} \quad (56)$$

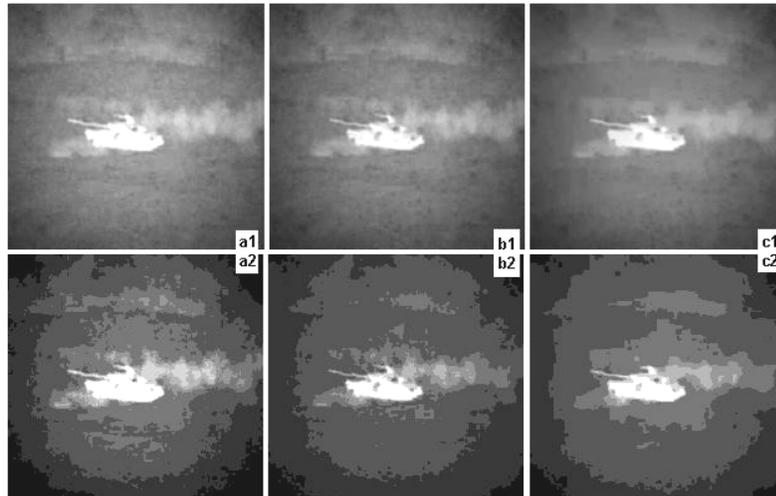
където: Y_{\max} и Y_{\min} са съответно максималната и минимална яркост в разглежданата област

Независимо че глобалният контраст е важен фактор, той не отговаря добре на начина, по който хората възприемат изображенията, особено в случаите, когато изображението изобилства от детайли. При тези случаи се използва дефинирането на локални контрасти за всяка една избрана зона. Съществената разлика на предложения тук метод с ползваните до момента (Stefan W., 2005), е че зоните в които се изчислява локалният контраст не са с фиксирани правоъгълни размери 5x5, 8x8 пиксела (стандартно за MPEG2), а се определят от ентропийните маски на изображенията. Глобалният контраст се изчислява като средно значение на всички локални контрасти на сегментираните обекти (фиг. 6.8):

$$\bar{C} = \frac{\sum_{i=0}^n C_i}{n} \quad (57)$$

където: \bar{C}_i е контраста на обектите, а n е броят на обектите.

Дефинирането на локалния контраст ни позволява да добием представа за детайлността на картината в зони 8x8 пиксела, обаче това не ни дава представа за общия контраст за отделните ентропийни канали.



Фиг. 6.8 Оригинално изображение – а), модово 3x3 рекурсивно филтрирано - b), медианно рекурсивно с тегла 3x3 - c)

Изчислението на глобалните контрасти за всеки един ентропиен канал се доближава до начина по който човек ще възприема определена област от изображението. Тези данни може да се ползват за дефиниране на каналите, които подлежат на допълнителна ентропийна сегментация, както и за изследване влиянието на различни филтри използвани за хомогенизация на изображението (таблица 6.2).

Контраст	Оригинално изображение	рекурсивна мода 3x3	Медиана 3x3
MSE		227	259
PSNR		56	55
Глобален	0.6653	0.6622	0.6303
канал 1	0.2115	0.2083	0.1959
канал 2	0.1141	0.1029	0.1029
канал 3	0.0995	0.1086	0.0824
канал 4	0.0568	0.0566	0.0686
канал 5	0.0388	0.0377	0.0303
канал 6	0.0386	0.0379	0.0317
канал 7	0.0563	0.0541	0.0667
канал 8	0.6653	0.6622	0.6303
локален	0.0435	0.0393	0.0243

Таблица 6.2. Сравнение на данните за контраста на изображенията от фиг. 22.

Изводи към Глава 6

- При анализа на видео филми средните приемливи отклонения за позицията на ключовите кадри са в рамките на 2-5 кадъра.
- Ползването на PSNR и MSE за метрика на качеството отразява абсолютното подобие между оригиналните и обработените последователности от изображения, като могат да бъдат прилагани и при анализа на серии от изображения.
- При детекция на глобални и локални ключови кадри се постига над 90% точност при сегментацията на видео филми, до 85% при сегментация на новинарски емисии с репортажни заставки и до 70% при спортни предавания (без анализ на аудио канала).
- Локалният контраст изчислен в рамките на ентропийно сегментираните маски, дава реалистична оценка на вижданото.

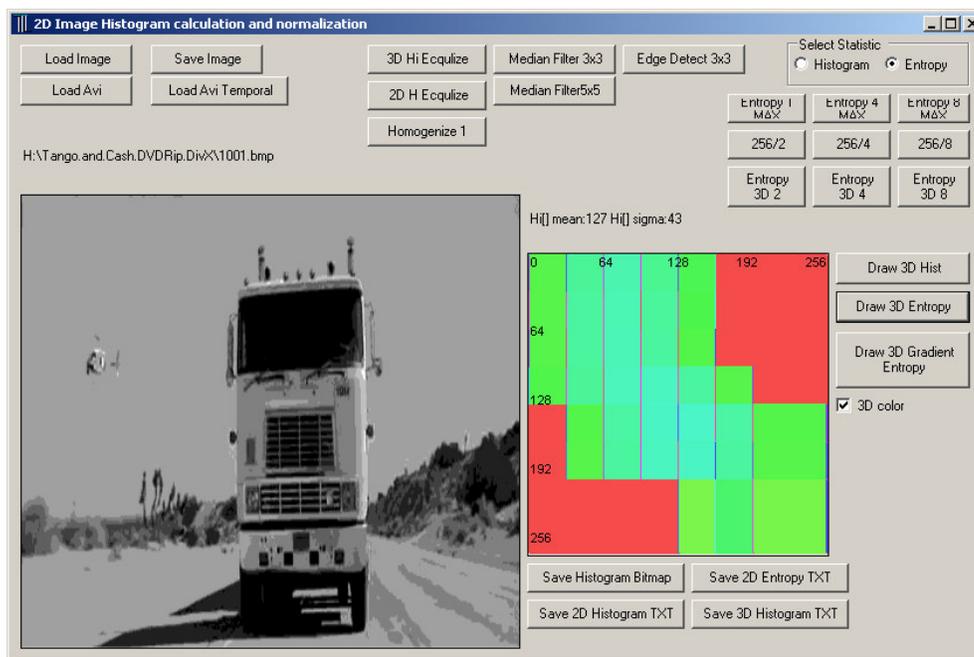
Приноси към Глава 6

- Разработен е и тестван алгоритъм за откриване на грешката при детекция на ключови кадри чрез изчисление на коефициентите **TKFC**, **MKFI** и **KFDI**.
- Предложен е модел за корекция на грешката при индексирание на мултимедийна информация, ползващ модата на цветните хистограми.
- Модифициран е методът за метрика на качеството с изчисление на глобалния контраст като сума от локалните контрасти в ентропийно сегментираните области.

Глава 7

СОФТУЕРНИ МОДУЛИ И ПРИЛОЖЕНИЯ ЗА АНАЛИЗ НА ПРЕДЛОЖЕНИТЕ В ДИСЕРТАЦИЯТА МЕТОДИ ЗА МНОГОМЕРНА ОБРАБОТКА И АНАЛИЗ

Цифровата обработка на последователности от изображения изисква наличието на бърз хардуер и реализацията на ефективни системи за обработка на видео филми в реално време върху стандартни компютърни архитектури е затруднено. Отчитайки тези особености, за нуждите на този дисертационен труд са разработени софтуерни приложения и филтри за обработка, сегментация и анализ на последователности от изображения, като показаното на (фиг. 7.3).



Фиг. 7.3. Софтуер за обработка и анализ на изображения

За получаване на експерименталните резултати са използвани и други програмни продукти като MathCAD и Excel (A6). За нуждите на този дисертационен труд са разработени следните оригинални авторски софтуерни филтри за обработка и анализ на видео информация (A4):

- Филтри за адаптивна филтрация на последователности от изображения;
- Филтри за 2D и 3D ентропийна сегментация на изображения;
- Модул за предварителна вероятностно статистическа сегментация на движението, откриване на ключови кадри и индексиране на визуални сцени, поддържащ автоматизираното разделяне на вече заснети видео филми в отделни файлове, ползващ експортиране на данните в MathCAD;
- Приложение за обектна сегментация, броене и класификация на обектите в отделни изображения;

Изводи към глава 7

- При имплементацията на софтуерни филтри преимуществено следва да се ползват 2D хистограмни и ентропийни модели за обработка поради своето бързодействие.

Приноси към Глава 7

- Реализирани са оригинални софтуерни приложения и филтри за обработка на видео в реално време, интегриращи се в софтуерния пакет с отворен код Virtual Dub.

Общи изводи от изследванията

Широкото навлизане на комерсиалните мултимедийни и интернет визуални приложения, системите за видео охрана и наблюдение изискват внедряването на автоматизирани интелигентни централизирани системи за управление и контрол.

В процесът на предварителна обработка и филтрация на изображенията се ползват алгоритми за 3D и 4D хистограмно групиране, адаптивна 2D и 3D модова или медианна филтрация.

Автоматизираното разделяне на видео филмите на отделни семантично свързани визуални сцени се извършва чрез вероятно статистическите критерии на Колмогоров, Пийрсън и Комплексен критерий и анализ на сегментираните цветни хистограми. В процесът на изчисление се анализират данните от 3D и 4D хистограмите и съответните ентропийни функции. Това позволява да бъде автоматизиран процеса на сегментация на визуални сцени.

Много параметричният векторен модел за представяне на обектите в изображенията позволява проследяване и параметрична обектна филтрация ускоряваща анализа на последователностите. С използването на векторите, описващи отделните изображения, и 4D хистограмите построява векторно пространство на движението, като самите ентропийно сегментирани 4D хистограми се ползват като идентификатори на визуалните сцени, позволявайки по този начин бързото сравняване и търсене на подобия между фрагментирани видео последователности. Векторното представяне на визуалните сцени би позволило създаването на формат за компресия с възможности за подобряване на компресията до 10 пъти в сравнение с DivX кодеците.

Разработен е и експериментално тестван модел за изчисление грешката при автоматизирано сегментиране на видео филми чрез индексите: TKFC, MKFI и KFDI. Използва се комплексна метрика на качеството, комбинираща PSNR и MSE и контрастите в локалните ентропийно сегментираните области.

Разработените софтуерни модули и приложения за обработка и анализ на последователности от изображения са тествани на следната компютърна система: CPU-AMD 1600MHz, 256MB RAM, както и Intel 2 Core Duo. Алгоритмите са мащабируеми, както по отношение на бързодействието така и по отношение на типа изходни последователности от изображения. Подобряване в скоростта на изчислението е възможно при ползване на по-бързи процесорни конфигурации.

Приноси на дисертационния труд

1. Създаден е метод за предварителна обработка на последователности от изображения чрез групиране на техните 3D хистограми, което позволява по-оптимално хомогенизиране на отделните яркостни преходи в изображението, което не може да се осъществи при групиране на 2D хистограмите. Оценени са възможностите за ползване на нелинейни 2D и 3D модови и медианни адаптивни софтуерни филтри за предварителната статистическа обработка на последователности от изображения. На база проведените експериментални изследвания е установено, че адаптивния 2D и 3D медианен филтър с тегловни коефициенти, дори при ползване на нецелочислени аритметични операции, позволява софтуерна реализация при максимални скорости на обработка на видео данните.
2. Предложена е модификация на комплексен вероятно статистически функционален модел ползван за детекция на движение в последователности от изображения, намиращ основно приложение в системите за видео охрана и наблюдение. Новост е ползването на последователностите от 3D хистограмите, което повишава чувствителността на алгоритъма при установяване на малки изменения, непостижими при ползване на 2D хистограмния модел. На тази основа е създаден адаптивен функционален модел, имащ и софтуерна реализация, позволяващ автоматизираната сегментация на видео филми чрез детекция на локални и глобални ключови кадри, които се ползват за индексване на начало и край на семантично свързани последователности от изображения обособени във визуални сцени. Използването на данните от сегментираните цветни хистограми за поредни изображения заедно със статистическите критерии на движението осигуряват по прецизна семантична разделимост на цифрови видео филми, което е непостижимо само при ползването на вероятно статистическите критерии ползвани за детекция на движението за нуждите на видео охранителните системи.
3. Разработен е метод за сегментация на изображения чрез многократно ентропийно групиране на техните 2D и 3D хистограми. На тази основа е създаден функционален модел за многоканална обектна сегментация на семантично свързани последователности от изображения. По този начин се постига допълнително минимизиране на случайната информация и шумове във видео последователностите преди процесът на обектна сегментация. Реализиран е много

параметричен векторен модел за представяне на отделните изображения и индексирание на визуални сцени. Този тип представяне на визуалната информация позволява извършването на параметрична филтрация и търсене на статистически представителни обекти с предварително зададени параметри в рамките на дадена визуална сцена.

4. Разработен е комбиниран 4D хистограмен модел за предварителна обработка, хомогенизиране и ентропиен анализ на последователности от изображения. Моделът предлага възможност за блокова обработка на сегментирани визуални сцени отчитайки и данните от обектната сегментация на отделните кадри. По този начин се осигурява възможност вече индексиранияте визуални сцени да бъдат сравнявани помежду си на база известни метрики за откриване на подобия в хистограми при силно редуциран набор изходни данни в сравнение с кадровото съпоставяне, което отнема много време, а също така на позволява възможност за съпоставяне на визуални сцени заснети с различна резолюция и скорост на кадрите 10, 15 или 30 кадъра в секунда.
5. Дефинирани са три критерия позволяващи оценка на грешката при автоматизирано детектиране на глобални ключови кадри. Извършено е експериментално изследване въз основа на което е предложен метод за корекция на грешката при автоматизираната сегментация на визуални сцени от видео филми чрез използване модата на цветните хистограми в съседни затварящи и отварящи визуалната сцена ключови кадри. Това минимизира грешката при неправилна фрагментация на видео последователности с високи коефициенти на вътрешно кадрово движение (каквато е случаят с детекция на движението при системите за видео охрана и наблюдение). Актуализирана е методологията за изчисление на метриката на качеството на изображения чрез пиковото отношение сигнал шум в яркостния канал на видео последователностите и чрез изчисление на глобалния контраст, като сума от локалните контрасти в ентропийно сегментираните области на изображенията.
6. Програмирани са авторски софтуерни филтри и отделни приложения за предварителна обработка, многомерен ентропиен анализ и сегментиране на последователности от изображения. Чрез тези софтуерни приложения е обезпечено успешното провеждане на експерименталните изследвания на представените в дисертацията методи и функционални модели за обработка и анализ.

Публикации по тема

- A1. Илиев П., Цветков П., **Петров Г.**, “Многомерный анализ динамических сцен проследивание движущихся объектов в последовательностях изображений”, International Scientific Journal of Computing, , Украина, 2007, том 6, выпуск 3, с. 30-37.
http://www.computingonline.net/issues/2007_vol_6_issue3.html#04
- A2. Илиев П., Цветков П., **Петров Г.**, “Многомерный анализ динамических сцен обнаруживание наличия движения с помощью анализа 3D гистограмм и 3D энтропийных функций”, International Scientific Journal of Computing, , Украина 2006Vol. 5, Issue 3, с. 35-43.
http://computingonline.net/issues/2007_vol_6_issue1.html#04
- A3. Илиев П., Tzvetkov P., **Petrov G.**, “Universal system for multi dimensional image analysis and recognition using Walsh – Hadamard functions”, сборник доклади от национален научен симпозиум с международно участие “Метрология и метрологично осигуряване” 2007, Созопол
- A4. **Петров Г.**, “Софтуерна компютърна видео охранителна система”, сборник доклади от национална конференция с международно участие “Телеком” 2006, Варна
- A5. Цветков П., **Петров Г.**, Илиев П., “Multidimensional dynamic scene analysis for video security applications”, сборник доклади от “Third International Bulgarian – Turkish Conference Computer Science’2006 Istanbul”
- A6. Цветков П., **Петров Г.**, Илиев П., “Software development environments for image acquisition and processing”, сборник доклади от международна научна конференция “Challenges in higher education and research in the 21st century”, Созопол 2006 – ISBN-10: 954-580-206-5
- A7. **Петров Г.**, Цветков П., Илиев П., “Детектиране на ключови кадри в цифрови видеозаписи”, сборник доклади от национален научен симпозиум с международно участие “Метрология и метрологично осигуряване” Созопол 2006 – ISBN 954-334-035-8
- A8. Цветков П., Илиев П., **Петров Г.**, “Цифрова обработка на изображения чрез използване на 3D энтропийна функция, анализ и класификация на микрообекти”, сборник доклади от национален научен симпозиум с международно участие “Метрология и метрологично осигуряване 2005” Созопол – ISBN 954-334-008-0
- A9. Илиев П., Tzvetkov P., **Petrov G.**, “Motion Detection Using 3D Image Histogram Sequence Analysis”, IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 5-7 September 2005, София, България ISBN 0-7803-9446-1, Library of Congress 2005931175, IEEE Catalog Number 05EX1192C
- A10. **Petrov G.**, “Low cost ECG system for non-hazardous use” IEEE International Spring Seminar on Electronics technology 27-Th. ISSE 2004 - ISBN 0-7803-8422-9
- A11. **Petrov G.**, Илиев П., Tzvetkov P., "Comparison of Global Histogram Methods for 2D and 3D Entropy Based Image Segmentation", p.57-62, 9th WSEAS International Conference on Evolutionary Computing (EC'08) ISBN: 978-960-6766-58-9, Sofia, Bulgaria, May 2-4, 2008