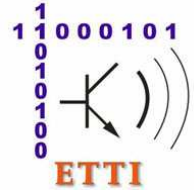




Universitatea “Politehnica” din București
Facultatea de Electronică, Telecomunicații și Tehnologia
Informației



Dezvoltarea unui Algoritm de Detecție Automată a Scenelor de Actiune din filme

-ghid de redactare a unui raport de cercetare-

Ș.l.dr.ing. Bogdan IONESCU¹

Coordonator:

Prof.dr.ing. Patrick LAMBERT
LISTIC – Laboratoire d’Informatique,
Systèmes, Traitement de l’Information
et de la Connaissance

2010

¹ acest manuscris reprezintă perspectiva autorului relativ la modul de alcătuire a unui raport de cercetare și a fost dezvoltat pentru a veni în ajutorul studenților. Datele prezentate sunt cu scop pur ilustrativ. De asemenea, în funcție de tema proiectului de cercetare, anumite componente sau informații propuse în acest ghid pot lipsi sau pot fi restructurate altfel.

Mulțumiri

- Această secțiune apare de regulă la lucrările de cercetare mai ample (lucrări de licență, teze de doctorat, etc.), totuși aceasta apare negreșit în momentul în care proiectul de cercetare este unul amplu și se realizează în cadrul unei instituții sau laborator de cercetare.
- În această secțiune se menționează persoanele care au avut un aport decisiv la buna desfășurare și finalizare a proiectului de cercetare, atât din punct de vedere științific cât și personal, precum profesorii îndrumători, alți cercetători, personal auxiliar, prieteni, etc.
- Tot în această secțiune sunt menționate părțile terțe care au pus la dispoziție resurse materiale (ex. finanțare studii, burse, ...) sau au furnizat accesul la infrastructura acestora de cercetare (ex. aparate de măsură, spații de lucru, ...).
- Această secțiune poate să lipsească în cazul în care proiectul de cercetare a fost realizat strict individual și cu resurse proprii, sau este vorba de un proiect restâns.
- *Doresc să mulțumesc cu această ocazie Dl. Profesor Patrick Lambert, Dl. Profesor Didier Coquin, Dl. Profesor Philippe Bolon și laboratorului LISTIC, Annecy-Franța, precum și Dl. Profesor Vasile Buzuloiu, Dl. Profesor Constantin Vertan, Dl. Profesor Mihai Ciuc și astfel laboratorului LAPI, București-România, pentru aportul lor decisiv la formarea mea profesională precum și pentru îndrumarea științifică prețioasă ce a permis indirect constituirea acestui ghid.²*

² pe întreg parcursul ghidului, exemplele propuse sunt scrise cu caractere italice pentru a ieși în evidență.

Cuprins

- Cuprinsul permite în primul rând o viziune de ansamblu asupra întregii lucrări, astfel acesta nu trebuie să lipsească. De asemenea acesta permite localizarea directă a părților din proiect ce prezintă interes pentru cititor:

1. Introducere	4
2. <i>Segmentarea temporală</i>	5
2.1. <i>Detecția tranzițiilor de tip cut</i>	5
2.1.1. <i>Metode bazate pe histogramă</i>	5
2.1.2. <i>Metode bazate pe contur</i>	5
...	
2.2. <i>Detecția tranzițiilor de tip fade</i>	5
2.2.1. <i>Metode bazate pe intensitatea pixelilor</i>	5
2.2.2. <i>Metode bazate pe contur</i>	5
...	
...	
5. Metoda propusă.....	8
6. Rezultate experimentale.....	9
7. Concluzii.....	11
8. Dezvoltări ulterioare.....	12
Bibliografie.....	13
Anexa 1.....	14
...	

1. Introducere

- Această secțiune obligatorie introduce tematica de cercetare familiarizând cititorul, într-o manieră concisă, cu domeniul în care se intervine.

- Informațiile care trebuie să transpară cititorului pot fi rezumate în:

- **obiectul cercetării:** se prezintă exact ceea ce se dorește să se realizeze, *exemplu: dezvoltarea unui algoritm de detecție automată a scenelor de acțiune din filme; și eventual ce direcție de studiu va fi folosită, ex. folosind analiza distribuției temporale a tranzițiilor video și a activității vizuale la nivel de plan video.*
- **scopul cercetării:** se scoate în evidență necesitatea acestei cercetări și impactul ei în domeniu, *exemplu: în prezent, informația multimedia, și în particular datele video joacă un rol important în viața noastră cotidiană (telefoane mobile, sisteme PC, sisteme de informare, Internet, corespondență, etc.). Accesul informației utile se dovedește a fi o problemă dificilă datorită volumului mare de date ce trebuie procesate de către utilizator. Soluția constă în indexarea automată a datelor. Sistemele existente sunt limitate în a folosi tehnici sintactice de analiză, în care accesul se realizează pe bază de informații matematice, numerice, greu accesibile persoanelor neavizate în domeniu. Proiectul propus vine cu o abordare nouă ce propune un nivel de analiză semantic, apropiat de modul de percepție uman, ...*
- **aplicabilitatea cercetării:** pe lângă aspectele științifice teoretice se pun în evidență aplicațiile practice directe și indirecte ale cercetării realizate, *exemplu: metoda propusă își găsește aplicație la eficientizarea managementul bazelor de date, la constituirea automată de rezumate de tip "trailer", clasificarea automată după gen a filmelor sau arhivarea filmelor...*
- **contextul cercetării:** ultimul aspect este legat de contextul practic al proiectului, aici menționându-se modul de concretizare, testare și validare practică a metodelor propuse: realizare în cadrul unui proiect de cercetare național/internațional, *exemplu: cercetarea a fost realizată în cadrul proiectului CNCSIS RP2 intitulat ..., parteneri ...* sau în cadrul unei întreprinderi private din domeniu, *exemplu: proiectul a fost realizat în colaborare cu CITIA, fiind testat pe baza de filme de animație din cadrul Festivalului Internațional al Filmului de Animație din Annecy, Franța.*

2. Segmentarea temporală

- Următoarele capitole, după cel introductiv, sunt dedicate principiilor teoretice ce stau la baza tematicii de cercetare abordată în proiect. În acestea este prezentat un studiu bibliografic al literaturii de specialitate din domeniu („state-of-the-art”) evidențiindu-se tehnicile existente precum și avantajele și limitările acestora.

- Scopul principal constă în familiarizarea cititorului cu aspectele teoretice din domeniu, precum și introducerea necesității cercetării realizate, prin identificarea absenței în literatura de specialitate a unei soluții viabile, sau prin identificarea unei posibile îmbunătățiri a unor soluții existente.

- Fiecare capitol poate fi structurat în mai multe subcapitole, subsubcapitole și așa mai departe, de exemplu:

2.1. *Detecția tranzițiilor de tip cut*

...

2.1.1. *Metode bazate pe histogramă.*

...

2.1.2. *Metode bazate pe contur.*

...

2.2. *Detecția tranzițiilor de tip fade*

...

2.2.1. *Metode bazate pe intensitatea pixelilor*

...

2.2.2. *Metode bazate pe contur*

...

- În cadrul redactării capitolelor se vor folosi următoarele informații:

- **referințe bibliografice:** prezentarea literaturii de specialitate existente se realizează prin descrierea și citarea lucrărilor existente: cărți, articole de revistă, articole de conferință, teze de doctorat, rapoarte de cercetare interne, resurse on-line, notițe de curs, etc.

Acestea sunt detaliate la capitolul de bibliografie și sunt referențiate prin intermediul unei etichete unice care poate fi un număr în paranteze pătrate, numele autorului urmat de anul publicării, etc., *exemplu: metodele de detecție a tranzițiilor de tip cut exploatează discontinuitatea vizuală produsă de acestea [1], discontinuitate ce poate fi măsurată folosind diferitele surse de informație ale secvenței: informația de contur [5], ... informația de mișcare [6], ... sau domeniul comprimat [7]. Metoda propusă în [2] realizează detecția pe baza ...*

Menționând sursa informațiilor prezentate se contribuie în primul rând la identificarea autorului acestora și astfel la conservarea drepturilor intelectuale, dar are și scopul de a furniza cititorului o referință la o documentație detaliată a subiectului prezentat succint în raport.

- **ecuații:** ecuațiile și formulele matematice sunt identificate de asemenea prin intermediul unei etichete de această dată exclusiv numerică și de regulă specificată între paranteze rotunde. Definirea unei relații matematice trebuie urmată de explicarea tuturor termenilor acesteia, exemplu:

... *histograma $h_I()$ a imaginii curente este dată de ecuația următoare:*

$$h_I(i) = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \delta(I(m, n) - i) \quad (1)$$

unde I este imaginea de dimensiune $M \times N$, i reprezintă nivelul de gri, $i=0, \dots, 255$ iar delta este dat de:

$$\delta(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases} \quad (2)$$

În text referințierea ecuațiilor se face pe baza numărului acestora, *exemplu: parametrul x depinde de histograma imaginii definită în ecuația 1.*

- **figuri:** figurile sunt folosite pentru a ilustra diverse principii, diagrame, imagini, rezultate, etc. Acestea sunt niște obiecte independente, flotante, ce pot apărea în diverse poziții în text, nu neapărat în locul în care sunt folosite. Fiecare figură conține în partea de jos o descriere ce explică conținutul acesteia și este identificată în text prin cuvântul Figura urmat de numărul acesteia, *exemplu: în Figura 1 am ilustrat diagrama metodei de detecție bazată pe histogramă sau pentru mai multe detalii vezi Figura 1.*

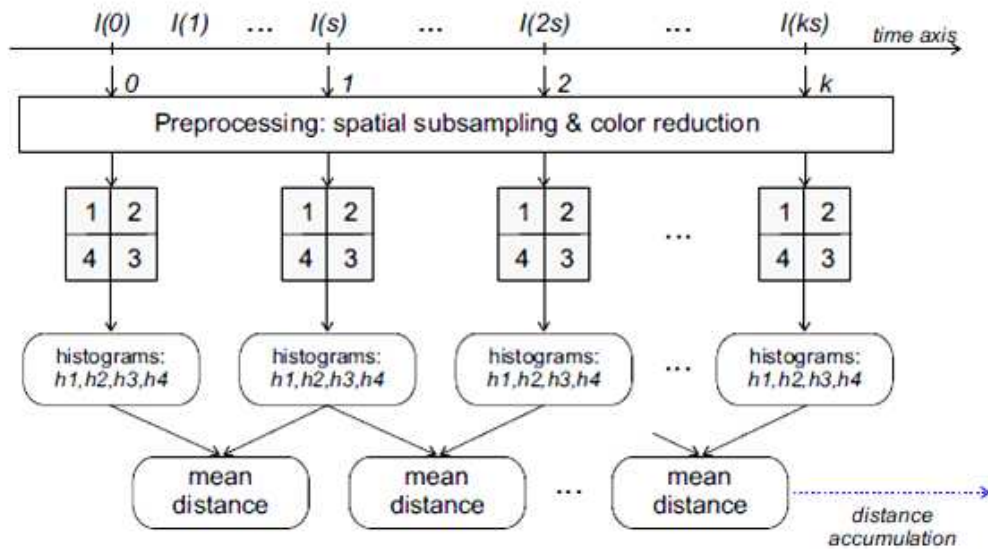


Figura 1. Diagrama metodei: $I(k*s)$ reprezintă imaginea la momentul de timp $k*s$, cu s pasul de analiză și k indicele temporal discret.

- **tabele:** tabelele păstrează aceeași convenție ca la figuri fiind precedate de această dată de descrierea conținutului acestora. Tabelele sunt referențiate folosind cuvântul Tabel urmat de numărul acesta, *exemplu: rezultatele testelor sunt prezentate în Tabelul 1:*

Tablelul 1: Comparație între metoda propusă și alte două abordări (GD=deteccii corecte, FD=deteccii eronate, SCC=schimbare scurtă de culoare).

Method	Count	GD	FD	Precision %	Recall %
conventional histogram-based	3166	2806	199	93.37	88.63
motion discontinuity-based	3166	2993	355	89.39	94.53
<i>proposed</i>	3166	2931	157	94.92	92.6
<i>proposed + SCC detection</i>	3166	2931	127	95.97	92.6

Pentru buna lizibilitate a conținutului unui tabel în antetul acestuia trebuie menționate unitățile de măsură ale mărimilor prezentate (dacă este cazul).

5. Metoda propusă

- Capitolul principal al raportului de cercetare este capitolul în care sunt descrise metodele sau tehnicile propuse. Acesta este realizat de regulă separat pentru a permite cititorului să identifice direct aportul autorului în domeniul abordat.

- Plecând de la principiul că o imagine face cât 1000 de cuvinte, se recomandă ca metodele să fie prezentate în principal folosind tehnici vizuale precum diagrame (vezi Figura 1 din Capitolul 2) și algoritmi, urmând apoi ca în cadrul capitolului să se descrie textual, în detaliu, fiecare etapă/componentă a metodei. De această dată trebuie ținut cont ca nivelul de detaliu furnizat de autor să permită cititorului să reconstituie integral, pe cât posibil, tehnicile propuse.

- exemplu: metoda propusă implică mai multe etape de prelucrare ce sunt detaliate în Figura 1. Astfel în primă fază secvența video este preprocesată fiind cuantizată temporal și spațial prin folosirea unui pas de analiză și prin reducerea rezoluției spațiale. Culorile sunt reduse folosind o cuantizare a spațiului de culoare RGB. Mai departe, fiecare imagine reținută este divizată în patru regiuni egale și pentru fiecare dintre acestea este calculată o histogramă color ... Etapele de prelucrare sunt detaliate în subcapitolele ce urmează ...

5.1. Preprocesare

...

5.2. Calcul histograme

...

5.3. Derivata de ordinul 2

...

5.4. Estimarea automată a pragului

...

6. Rezultate experimentale

- La fel de important precum capitolul ce descrie principal tehnicile propuse este capitolul în care sunt prezentate rezultatele practice obținute în urma testării acestora într-un context concret, sau cu alte cuvinte validarea practică a metodelor. Acest capitol este definitoriu pentru credibilitatea cercetării realizate și pentru justificarea aportului semnificativ adus de aceasta în domeniu.

- Acest capitol, în mod ideal trebuie să puncteze următoarele aspecte:

- **modul de testare:** trebuie specificat modul în care au fost testate metodele și sursa datelor care au fost folosite pentru aceasta, ex. bază de date de test etichetată ("groundtruth"), realizarea unor studii individuale, etc. Pentru a conferi încredere metodei, aceasta trebuie testată pe un eșantion suficient de vast de exemple și în condiții suficient de generale, *exemplu: metoda propusă de detecție de cuts a fost testată pe o bază de filme de animație ce conține 100 de filme cu tehnici variante (ex. animație 3D, animație de obiecte, desen pe hârtie, etc.), cu o durată totală de 10 de ore și 40 minunte și un număr de 15830 de tranziții de tip cut (vezi Anexa 1). Acestea au fost etichetate manual pentru a servi ca groundtruth pentru validarea metodei propuse.*
- **modul de apreciere al performanțelor:** pe lângă datele folosite la testare trebuie precizat modul în care a fost evaluată performanța metodelor sau tehnicilor propuse, ceea ce poate fi realizat subiectiv pe baza expertizei umane (ex. grad de satisfacție) sau numeric prin furnizarea unor măsuri matematice ce cuantifică erorile metodei, exemplu:

... pentru evaluarea performanței metodei propuse s-au folosit erorile clasice de tip "precision" și "recall" ce sunt definite în ecuația 3:

$$precision = \frac{GD}{GD + FD}, recall = \frac{GD}{N} \quad (3)$$

unde GD reprezintă numărul de detecții corecte, FD reprezintă numărul de detecții eronate și N reprezintă numărul total de tranziții existente.

- **testarea propriu-zisă și compararea cu alte metode:** sunt prezentate rezultatele obținute de metodele propuse în contextul menționat anterior și pentru diverse configurații ale parametrilor de reglaj, evidențiindu-se setul de parametri optimi ce conduce la rezultatele cele mai bune (ex. eroare minimă). De asemenea, sunt puse în evidență și comentate situațiile în care metoda furnizează cele mai bune rezultate, identificând avantajele acesteia, dar și o serie de exemple în care aceasta furnizează rezultate slabe sau chiar eșuează, explicând producerea acestora și identificând astfel potențialele limitări.

Un aspect foarte important ce conferă credibilitate metodei este compararea acesteia cu alte metode existente, testate în aceleași condiții. Se caută să se dovedească superioritatea tehnicilor propuse și astfel utilitatea și gradul de noutate al acestora, *exemplu: în Tabelul 1 metoda propusă a fost comparată cu două metode clasice, o primă metodă bazată pe distanța între histograme (notată histogram-based) și o metodă bazată pe discontinuitatea câmpului de mișcare (notată motion discontinuity-based). Metoda propusă furnizează atât o rată de detecție eronată scăzută (precision >95%) cât și un număr de non-deteccții scăzut (recall >92), furnizând rata de precision și recall cea mai bună.*

Modul de prezentare al rezultatelor experimentale trebuie să fie unul cât mai vizual, bazat pe imagini, grafice, tabele comparative, etc. întărite de comentarii și interpretarea conținutului lor.

7. Concluzii

- În acest capitol sunt prezentate concluziile generale asupra cercetării propuse și a testelor practice. De regulă la început se recapitulează foarte concis (unul sau două paragrafe) ceea ce s-a realizat, furnizând și câteva detalii asupra tehnicilor folosite (similar introducerii).
- Mai departe se trece la modul de testare experimentală și la concluzionarea rezultatelor experimentale evidențiind performanțele metodelor sau tehnicilor, precum și experiența personală a autorului dobândită în urma experimentării practice (dificultăți, limitări, etc.). Tot aici sunt evidențiate posibilele puncte slabe și limitări ale metodei sau tehnicilor dezvoltate.
- De regulă acest capitol nu depășește o pagină.

8. Dezvoltări ulterioare

- Acest capitol conține de regulă câteva paragrafe în care este explicat modul de continuare a cercetării propuse în raport, prin extinderea metodelor propuse, abordarea limitărilor acestora, creșterea robusteții, abordarea unor direcții noi, etc., *exemplu: posibile îmbunătățiri ale metodei de detecție de cut propusă constau în principal în fuziunea mai multor modalități ale secvenței: intensitate-mișcare-contur, pentru a beneficia astfel de avantajele furnizate de fiecare dintre ele.*

Bibliografie

- În această secțiune sunt menționate toate materialele ce au contribuit la dezvoltarea proiectului de cercetare, și anume: cărți, articole de revistă, articole de conferință, teze de doctorat, rapoarte de cercetare interne, adrese către resurse on-line, notițe de curs, etc.

- Fiecare material este precedat de o etichetă unică pe baza căreia este citat în text, acolo unde intervine sau sunt folosite informații din acesta.

- Exemplu de referințe bibliografice:

[1] B. Ionescu: „*Analiza și Prelucrarea Secvențelor de Imagini: Indexarea Automată după Conținut*”, Editura Tehnică București, ISBN 978-973-31-2354-5, 2009.

[2] B. Ionescu, D. Coquin, P. Lambert, V. Buzuloiu: „*A Fuzzy Color-Based Approach for Understanding Animated Movies Content in the Indexing Task*”, Eurasip Journal on Image and Video Processing, special issue on Color in Image and Video Processing, doi:10.1155/2008/849625, 2008.

[3] B. Ionescu, A. Păcureanu, P. Lambert, C. Vertan, "*Highlighting Action Content in Animated Movies*", IEEE ISSCS - International Symposium on Signals, Circuits and Systems, special session on Action Recognition and Tracking in Multimodal Video, 9-10 Iulie, Iași, Romania, 2009.

[4] B. Ionescu: „*Content-Based Semantic Retrieval of Video Documents, Application to Navigation, Research and Automatic Content Abstraction*”, proiect CNCSIS RP-2, <http://alpha.imag.pub.ro/VideoIndexingRP2/>.

[5] Z. Huan, L. Xiuhuan, Y. Lilei: “*Shot boundary detection based on mutual information and canny edge detector*”, IEEE International Conference on Computer Science and Software Engineering, pag. 1124-1128, 2008.

[6] S.V. Porter, M. Mirmehdi, B.T. Thomas: “*Video cut detection using frequency domain correlation*”, International Conference on Pattern Recognition, pag. 413–416, Barcelona, Spain, 2000.

[7] J. Jiang, Z. Li, G. Xiao, J. Chen: “*Real-Time Shot-Cut Detection in a Compressed Domain*”, SPIE Journal of Electronic Imaging, 16(4), 2007.

Anexa 1 – *Baza de filme folosită*

- Anexele sunt ultimele secțiuni ale raportului și sunt facultative. Acestea includ de regulă: materiale facultative, explicitarea unor tehnici clasice din domeniu permițând astfel cititorului neavizat să se familiarizeze cu acestea, date voluminoase ce nu pot fi prezentate în text precum imagini, tabele, codul sursă al unor programe, etc. Un raport de cercetare poate conține mai multe astfel de anexe ce sunt identificate prin numărul acestora ca orice capitol din raport.