



Universitatea Politehnica din București

TEZĂ DE ABILITARE

**Abordări de inspirație perceptuală în
prelucrarea și analiza imaginilor**

Corneliu Florea

Facultatea de Electronică, Telecomunicații și Tehnologia Informației

Departamentul de Electronică Aplicată și Ingineria Informației

Laboratorul de Analiza și Prelucrarea Imaginilor

1 REZUMAT

Prezenta teză de abilitare are drept obiectiv oferirea unei imagini de ansamblu asupra evoluției carierei academice a candidatului, a principalelor sale realizări științifice și să ofere o perspectivă asupra activității viitoare. Lucrarea reflectă o parte din direcțiile și preocupările recente, de după obținerea titlului de doctor inginer în 2009. Un scurt rezumat este prezentat în capitolul introductiv.

Tema **Capitolului 2**, care este totodată și primul capitol tehnic, o constituie studiul modelelor de tip logaritmice pentru prelucrarea imaginilor. Dacă modelul logaritmice de prelucrare a imaginilor există din 1987, în ultima perioadă au apărut noi extensii. Astfel, în 2008 a apărut un nou model de tip pseudo-logaritmice. În lucrarea de față sunt descrise contribuțiile recente ale candidatului la dezvoltarea acestui model. În primul rând se introduce un model de parametrizare prin compunerea funcției generatoare cu o funcție parametrică. Acest model parametric oferă o mai mare flexibilitate, care este exploatată în construcția unei noi funcții de estimare a clarității unei imagini (in-focus). Curbarea extremă pentru valori mici ale intensității luminoase permite performanță sporită în cazul imaginilor sub-expuse.

Cea de a doua contribuție discutată în acest capitol se referă la identificarea similarității cu percepția umană a modelului de tip logaritmice. Mai precis se arată că modelul este cvasi-similar cu modelul Naka-Rushton al percepției umane. Utilizarea acestei similarități a fost folosită pentru construcția unei metode de extindere a gamei dinamice (HDR) a cărei principală calitate este de a permite o abordare unitară pentru cele două alternative mari: fuziunea cadrelor, respectiv fuziunea hărților de radianță. În paralel, modelul nou creat a condus la rezultate, care, per-ansamblu, sunt mai bune decât ambele variante existente.

Capitolul 3 se concentrază pe analiza feței umane și descrie o metodă pentru identificarea intensității durerii pe baza scorului Prkachin-Solomon. Tehnic, contribuția inovatoare descrisă în acest capitol este duală. În primul rând sunt introduse histogramele de trăsături topografice care utilizează și informații extrase din matricea Hessiană locală a fiecărui pixel. A doua contribuție se referă la introducerea unei noi metode de transfer learning (învățare prin transfer) în care numărul mic de persoane dintr-o bază de date medicală este compensat prin utilizarea unei baze extinse, dar cu altă tematică, pentru identificarea celor mai puternice dimensiuni descriptive. Sistemul este competat de un ansamblu format din mașini cu vectori suport pentru clasificarea propriu-zisă. Performanța pe secvențe este crescută prin exploatarea dimensiunii temporale și a intercorelației dintre cadre. Dacă în capitolul anterior omul era referința pentru inovația tehnologică, în cazul de față, omul devine subiectul studiului.

Capitolul 4 se concentrează din nou pe componenta umană. De data aceasta tema centrală este arta și se prezintă două sisteme care au pornit pe drumul de a afla dacă putem înțelege, în mod automat, ceva atât de specific uman precum este arta. În acest capitol se descriu două contribuții pentru recunoașterea automată a curentului artistic.

Pentru prima soluție discutată descrie un sistem ce imită extensiv percepția umană de nivel mediu. Dacă energia filtrelor Gabor a mai fost folosită pentru descrierea artei, volumul color dominant, respectiv informații asociate șabloanelor extrase prin teoria ancorării vizuale constituie contribuții inovative. Pentru a efectua recunoașterea efectivă a curentului artistic se utilizează o schemă de fuziune târzie bazată pe combinarea unui nivel alcătuit de 3 perceptoare multi-strat cu un nivel agregator bazat pe o mașină cu vectori suport.

A doua soluție își propune în primul rând o evaluare mai precisă a performanței. Acest lucru este realizat prin introducerea unei baze de date de volum mare alcătuită din imagini ale tablourilor artistice; baza de date a trecut inițial printr-un proces atent de selecție și adnotare. În al doilea rând, se introduce un nou sistem de clasificare automată, în care descrierea imaginilor este duală: prin structura de culoare respectiv prin textură. Ultima este realizată cu ajutorul caracteristicilor topografice care își dovedesc și în acest caz utilitatea. Clasificarea este realizată de către un ansamblu amplificat (boosted) de mașini cu vectori suport. Sistemul reușește să izoleze datele clasificate inițial incorect, iar pentru a corecta aceste erori, le crește iterativ importanța. Performanța rezultată a sistemului este superioară unor alternative clasice din literatură.

În **capitolul 5** tema abordată este cea a localizării și recunoașterii logourilor în imagini naturale. Problematika localizării și recunoașterii logourilor este dificilă în imagini naturale, datorită deformărilor de perspectivă, variabilității fundalului, a posibilităților de ocluzii, sau scalării variabile; în plus, campaniile de re-branding induc schimbări în paleta de culori și în distribuția spațială a unui logo. Toate aceste aspecte creează dificultăți unui sistem automat, dar sunt rezolvate instantaneu de către un om. Soluția propusă presupune localizarea punctelor cheie folosind descriptorul afin al Diferențelor Gaussiene. Vecinătatea acestora este aleasă a fi eliptică pentru a putea identifica mai precis orientarea logoului. Fiecare marcă este descrisă de o clasă iar metoda presupune construcția de prototipuri de clasă prin analiza graficului de potrivire între exemple omografice. Graficul de interconexiuni este dezvoltat pentru fiecare clasă și se adaugă punctele reprezentative din exemplele non-centrale la modelul de clasă. Fiecare clasă are un potențial model secundar care este inversat; decizia de a-l construi este luată în mod automat. În final, fiecare clasă este reprezentată de reuniunea corespunzătoare de puncte cheie și descriptori. Recunoașterea clasei și detecția sunt integrate. Performanța obținută indică cele mai bune rezultate comparativ cu lucrările anterioare din literatură pe aceeași temă, pe trei baze de date publice.

Ce a de a doua parte, și totodată **capitolul 6** a lucrării se concentrează pe aspecte administrative. Se prezintă principalele realizări academice ale candidatului cu accent pe experiența acumulată în ultimii ani. O dată realizat sumarul stadiului actual se identifică direcții de urmărit în evoluția viitoare, atât din punct de vedere al cercetării cât și administrative. Drept principale puncte forte sunt identificate numărul mare de publicații din ultima perioadă, legăturile cu industria și colaborarea strânsă cu foștii studenți.

Carieră universitară a candidatului, de la debut până în prezent, s-a desfășurat în cadrul Laboratorului de Analiza și Prelucrarea Imaginilor. Colegilor mei din acest laborator le sunt profund recunoscător pentru permanentul sprijin și sfaturile oferite. Familia mea a fost mereu un element de susținere pe parcursul tuturor acestor ani, fapt pentru care trebuie să le mulțumesc.