

IMPLEMENTAREA SISTEMELOR BIOMETRICE AUTONOME PENTRU RECUNOAȘTEREA FEȚEI UTILIZÂND PROCESOARE DSP MEDIA

Radu Arsinte

Communications Department, Technical University, Cluj-Napoca, 26-28 Gh. Barițiu Street

Abstract –The paper presents a study of autonomous face recognition systems based on high performance DSP, so called Media Processors. An overview of the most powerful Media Processors available today is followed by an analysis of the common features of the face recognition algorithms used to implement an application. The paper presents also a generic system implemented using DSP Media processors, both from a hardware or software perspective. Results and benchmarks for the generic system are also presented at the end of the paper.

Key words: Biometry, Face recognition, Video capture, DSP, Media processors

INTRODUCERE

Sistemele biometrice și supravegherea video electronică sunt două domenii în plină dezvoltare care atrag interesul atât al industriei cât și al comunității științifice. O aplicație importantă în aceste sisteme este recunoașterea feței, care constă în identificarea unei persoane pe baza imaginii feței sale. Deși domeniul de cercetare este vechi de peste 30 de ani, o activitate mai intensă se observă după anul 1990. Aceasta a dus la dezvoltarea de algoritmi eficienți și de introducerea de produse comerciale bazate pe acești algoritmi. Prin trecerea algoritmilor din laboratoare în aplicații concrete, factori aparent secundari, cum sunt consumul de putere și costul sistemului, au devenit esențiali. Aceste aspecte au dus în mod natural către implementări bazate pe DSP ca o platformă eficientă. O categorie aparte de DSP, cu capabilități excepționale în domeniul prelucrării video sunt procesoarele media descrise în [1, 2]. Ele sunt utilizate cu succes în implementarea unor sisteme biometrice autonome (*embedded* după denumirea din limba engleză)

Procesoare media

Sistemele bazate pe procesoare media au la dispoziție mai multe platforme care pot fi utilizate în implementare. Dintre acestea am ales pentru exemplificare două familii de procesoare care prin maturitate și suportul oferit de firme puternice și cu tradiție în domeniu sunt interesante în mod particular : Texas Instruments TMS320C64xx și Philips Nexperia.

TMS320C64xx - Texas Instruments

Generația de procesoare TMS320DM64x oferă codare , decodare și transcodare multicanal pentru toate standardele actuale în domeniul video și audio. Sunt dispozitive complet programabile care pot fi utilizate în toate aplicațiile care necesită prelucrare complexă în domeniul video și audio. Toate acestea, prin complexitatea circuitului, revin la un preț foarte redus.

Câteva caracteristici:

- Porturi video integrate de înaltă rezoluție ;

- Posibilitatea de prelucrare simultană a până la patru fluxuri MPEG-2 (profil/nivel bază) cu decodare în timp real;
- Codare MPEG-2 în timp real
- Suport pentru Windows Media Encode 9
- Conectivitate Ethernet și PCI integrată

Nexperia- Philips

PNX-1300 este un circuit programabil cu o construcție ideală pentru aplicații care necesită prelucrarea simultană a tipurilor de date din fluxurile multimedia. Cu o suficientă putere disponibilă de calcul pentru a captura, comprima și decompresa multe formate de date video și audio în timp real, PNX-1300 este bine adaptat pentru un domeniu larg al aplicațiilor centrate pe prelucrarea video ca: videoconferință, editare video, baze de date video, control, sau sisteme industriale de inspecție și circuite multifuncționale ca receptoare TV digitale și sisteme de supraveghere.

Caracteristici

- Datele, audio, video, grafice și fluxuri de date se prelucrează pe un un singur circuit;
- Frecvența de funcționare 143 sau 166 MHz, capacitatea de prelucrare a VLIW CPU ajungând până la 6,5 BOPS;
- instrucțiuni multiple: de tip microprocesor tradițional, SIMD multimedia și operații cu virgulă flotantă format IEEE;
- Este posibilă dezvoltarea de aplicații în limbajele de programare C și C++;
- Biblioteci de prelucrări elementare (video, imagini) puse la dispoziție de Philips.

Complexitatea circuitelor din familia Nexperia este ilustrată de figura 1 (preluată din [2]).

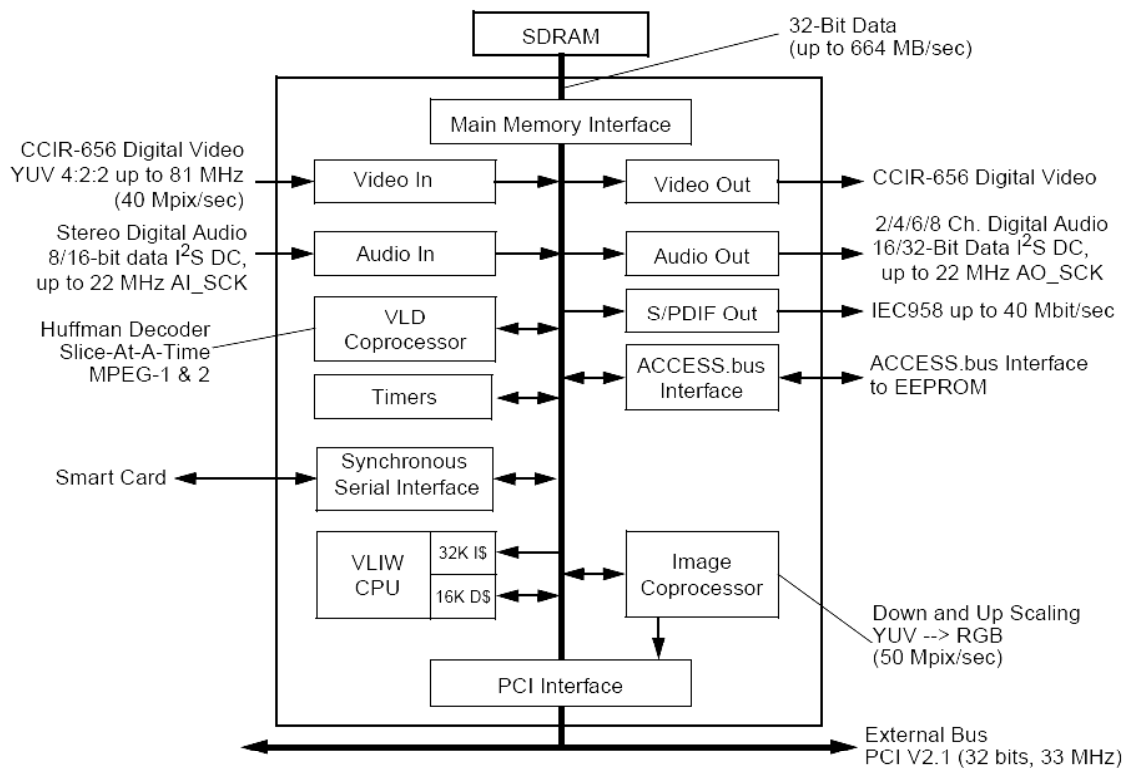


Figura 1. Structura procesoarelor Nexperia PNX1300

ALGORITMI UTILIZAȚI ÎN RECUNOAȘTEREA FEȚEI

Scopul recunoașterii feței este de a determina identitatea unui personaj bazându-ne pe o imagine statică sau într-o secvență video a feței sale. Sunt două moduri de funcționare pentru un sistem de recunoaștere a feței: autentificare și identificare. În modul de autentificare sistemul acceptă sau respinge identitatea revendicată de individ. În modul de identificare sistemul compară fața cu o bază de date a unor persoane cunoscute și returnează cea mai probabilă identitate sau identități. În funcție de sursa de imagine pentru identificare, algoritmul de recunoaștere are abordări diferite fiecare cu avantajele și dezavantajele sale.

Recunoașterea fețelor poate fi încadrată în contextul mai larg al metodelor de recunoaștere a imaginilor [3]. În această abordare o metodă tipică de recunoaștere are trei etape (figura 2):

- Transformarea imaginii primare într-o reprezentare inițială (poate fi reprezentată de o preprocesare sau de o transformare matematică, de exemplu calculul valorilor proprii);
- Extragerea caracteristicilor cheie (păstrarea primelor n componente pentru DCT, de exemplu);
- Mecanismul de clasificare (modelare): modele bazate pe clusteri, metrice, rețele neuronale, etc.

Astfel, construcția metodei de recunoaștere se bazează pe informații apriorice despre domeniul investigat (în acest caz imaginea feței umane), și corectarea cu informație experimentală, apărută în cursul elaborării metodei.

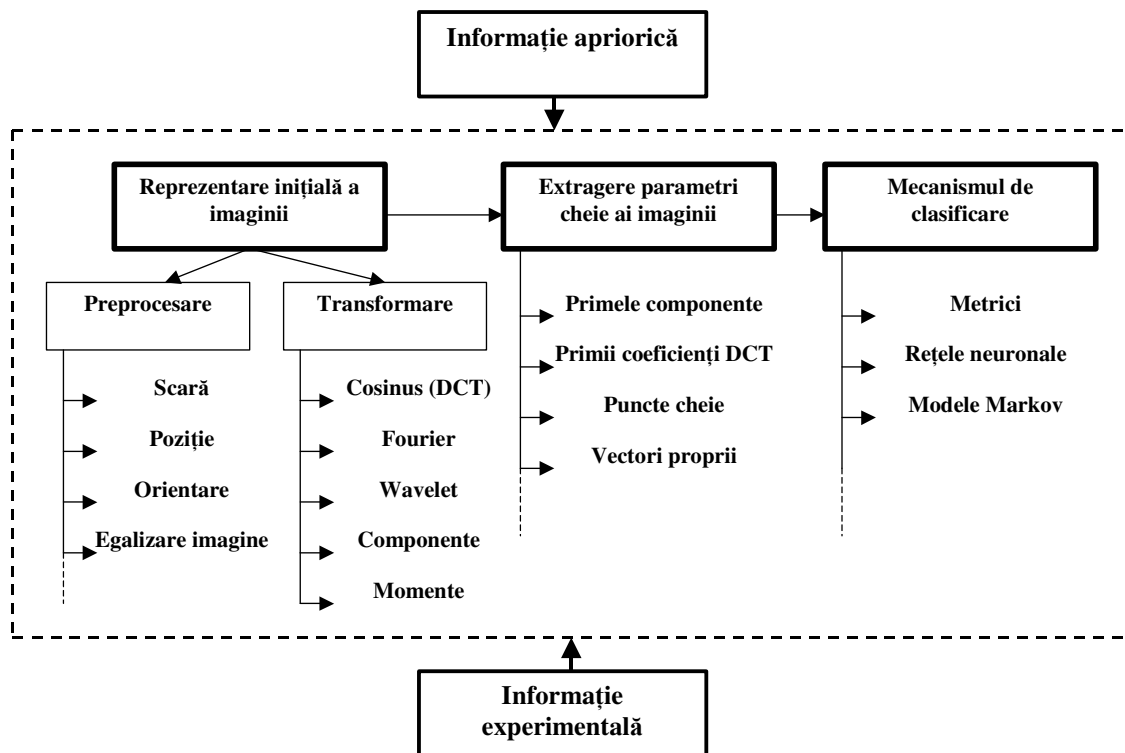


Figura 2 – Structura metodelor de recunoaștere a imaginilor

Figura 3 prezintă schemele bloc pentru cele două abordări ale recunoașterii feței. Pentru simplitate se presupune că avem o singură față în imagine sau secvență video. În cazul existenței a mai multor fețe în imagine sistemul lucrează separat pe fiecare din acestea. În diagramele din figura 3 trebuie menționat faptul că căutarea și extragerea caracteristicilor sunt taskuri care în mod particular necesită calcule intensive.

Algoritmii de recunoaștere analizați în [6] folosesc abordarea „învățării” vizuale propusă de Moghaddam and Pentland [4]. În această abordare, imaginile feței sunt modelate ca o distribuție Gaussiană care este estimată cu reducere dimensională folosind transformata Karhunen-Loewe (KLT). Pentru a detecta fețe în imagini statice, blocuri la scări diferite și localizate în diverse locuri sunt extrase din imagine, și se calculează probabilitatea de a fi o față folosind densitatea de probabilitate menționată în [6].

Deoarece căutarea unei imagini mari la scări multiple și în locații multiple est un task computațional intens, se încearcă reducerea spațiului de căutare pe baza unei reguli propuse în [6]. Presupunând că avem doar o față în imagine, metoda sugerează că modificările abrupte în profilele verticale sau orizontale ale imaginii corespund cu limitele capului. Profilul orizontal este obținut prin medierea pixelilor pe fiecare coloană, iar profilul vertical este obținut prin medierea pixelilor pe fiecare linie. Totuși, în cazul general, dacă o persoană este situată pe un fundal oarecare este puțin probabil să găsim limitele capului prin această metodă, deoarece pot fi și alte schimbări în fundalul imaginii care să producă același efect. În acest caz regula este utilă în reducerea spațiului de căutare al feței. În profilul orizontal, se determină prima și ultima modificare abruptă și se presupune că cel puțin jumătate din față este situată între aceste două limite. Dacă sunt detecții false datorate fundalului, aceasta mărește spațiul de căutare fără a pierde fața propriu-zisă. În mod similar, în profilul vertical, se găsește prima modificare abruptă, și se presupune că fața este situată sub limita superioară. În spațiul dintre aceste limite se face o căutare prin metoda propusă de Moghaddam și Pentland în locuri

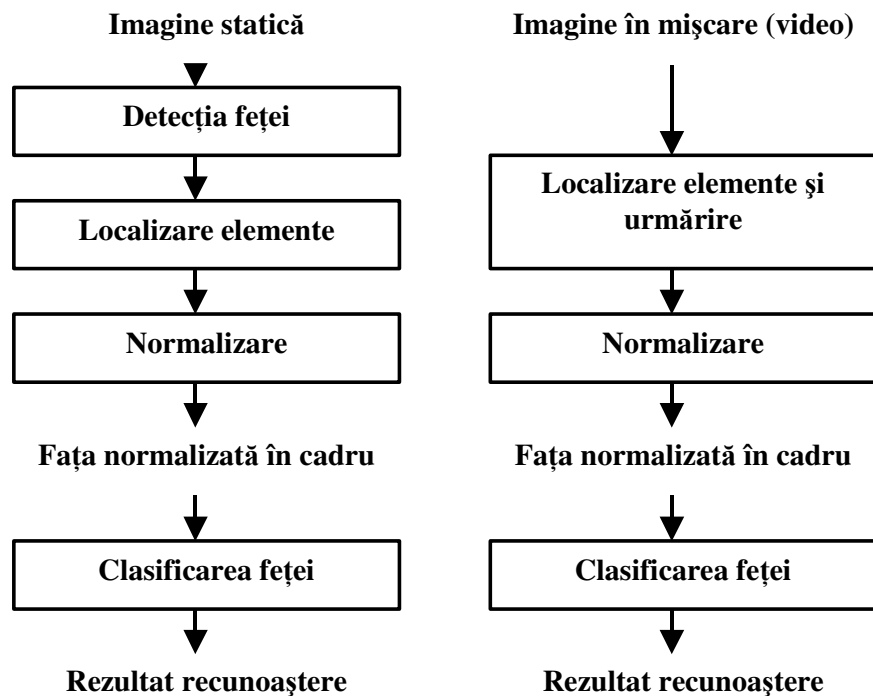


Figura 3. Variante ale algoritmilor de recunoaștere a feței

diferite și la scări multiple așa cum s-a explicat anterior. După găsirea poziției feței, se face căutarea poziției ochilor, prin căutarea celor doi ochi de asemenea la scări multiple și în locuri diferite. Se utilizează estimarea de densitate propusă de Moghaddam și Pentland pentru modelarea distribuției ochilor. După găsirea ochilor, imaginea este rotită pentru a avea ochii în poziție orizontală, este mascată pentru a exclude fundalul și este decimată pentru reducerea ei la o rezoluție de 128x128 pixeli.

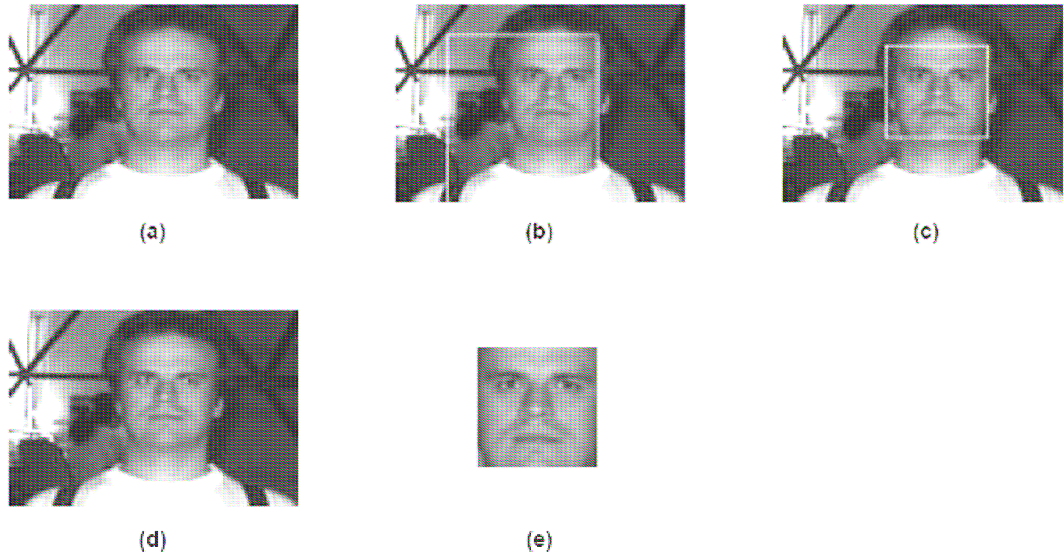


Figura 4. Detecția feței, localizarea ochilor și etapele de normalizare a fețelor (a) imagine originală, (b) reducerea spațiului de căutare, (c) rezultatul căutării feței, (d) rezultatul căutării ochilor, (e) spațiul feței normalizat

Acești pași se numesc normalizarea feței. Ei sunt ilustrați de figurile 4 și 5.

După ce imaginea feței este normalizată, este trimisă algoritmului de clasificare a feței care o compară cu baza de date a persoanelor cunoscute și returnează cea mai probabilă identitate. Se prezintă două versiuni ale algoritmului de clasificare a fețelor. Primul este mai cunoscut sub numele de „Eigenface” propus de [4], care este algoritmul de bază în recunoașterea fețelor. În această abordare, imaginile sunt proiectate într-un subspațiu care este obținut prin realizarea analizei componentelor principale pe imaginile de antrenament. Recunoașterea este realizată prin clasificarea pe baza distanței minime.

A doua metodă de clasificare este a subspațiilor liniare segmentate, algoritmul propus de Batur și Hayes [5]. Scopul acestui algoritmul este de a realiza o recunoaștere fiabilă în condiții de iluminare diverse. În această abordare, imaginea feței fiecărei persoane cu o poză impusă achiziționată în condiții variate de iluminare, este modelată cu un model spațial segmentat liniar, iar recunoașterea este realizată prin calculul distanței între imagine și modelele subspațiale din baza de date. Procedurile de recunoaștere sunt implementate inițial într-un limbaj adecvat, de exemplu Matlab, urmează testarea acestora și apoi validarea algoritmului. În exemplul prezentat [6] baza de date de validare conține 300 fețe, în imagine frontală, a 10 persoane în care direcția de iluminare variază între 10 și 50 de grade.

Pentru sistemul automat rata de recunoaștere pentru metoda „eigenface” a fost de 88% , iar pentru subspațiile segmentate liniar de 93%.

Aceste metode au fost implementate în sistemul bazat pe TMS320C64xx.



Figura 5. Detecția feței, localizarea ochilor și normalizarea feței: (a) imaginea originală, (b) reducerea spațiului de căutare, (c) rezultatul al etapei de detecție a feței, (d) rezultatul etapei de localizare a ochilor, (e) imaginea feței normalizate

IMPLEMENTAREA SISTEMELOR DE RECUNOAȘTERE A FEȚELOR

Structura sistemului

Această secțiune prezintă două sisteme generice de recunoaștere a fețelor prezentate în [6] și [7]. Ele sunt bazate pe procesoarele media descrise în paragraful introductiv.

Implementarea sistemelor de recunoaștere a feței poate fi făcută pe platforme de uz general, de tipul calculatoarelor compatibile PC, sau în sisteme autonome care minimizează consumul, dimensiunile și costurile. Tendința generală este către „autonomizarea” sistemelor care permit instalarea acestora în locurile cele mai potrivite pentru un control al identității eficient (aeroporturi, gări, porți de acces).

În continuare vom prezenta un sistem generic [7] de control biometric bazat, așa cum am argumentat în paragrafele precedente, pe un DSP performant.

Datele tehnice ale sistemului generic de recunoaștere a feței sunt prezentate sintetic în tabelul 1. Flexibilitatea sistemului cu DSP permite ca pe aceeași structură hardware să poate fi abordate aplicații variate din domeniul de prelucrare video : biometrie, viziune industrială, supraveghere video.

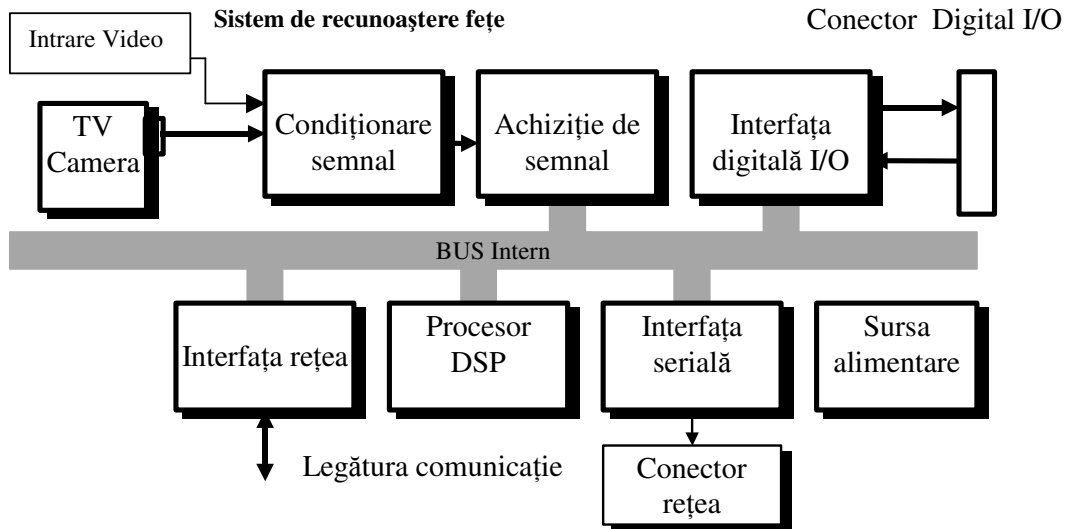


Figura 6. Structura unui sistem de recunoaștere a feței bazat pe DSP media

Rezoluție achiziție	Minimum 640x480 (768x576) pixeli
Semnal intrare	Analog – PAL/SECAM compatibile Digital – semnale LVDS
Interfața comunicație	Ethernet 100 MBit/s
Memorie (Flash)	4-8 MB
Memorie (RAM)	16-64 MB
Intrări/ieșiri auxiliare	8 intrări, 8 ieșiri
Funcții speciale	Vizualizare semnal direct (comprimat / necomprimat)
Facilități opționale	Interfața digitală auxiliară (USB, IEEE1394)

Tabel 1

Principalele blocuri evidențiate în figura 2 au următoarele funcții.

Blocul de condiționare a semnalului realizează în principal multiplexarea semnalelor analogice, amplificarea acestora, funcții de sincronizare și restaurare semnal video.

Partea de **achiziție și normalizare** este implicată în conversia A/D, formatarea semnalului în standardele acceptate de interfața video a DSP, normalizarea imaginii.

Unitatea centrală cu DSP are controlul tuturor funcțiilor de achiziție și de recunoaștere incluse în aplicație, controlul comunicației, detectarea pragurilor și decizii legate de aceasta, stocarea/refacerea informației.

Arhitectura software

Partea de suport software a sistemului este formată în principal din două module: nucleul sistem de operare (preferabil de timp real) și aplicația propriu-zisă.

Nucleul sistemului de operare este impus în general de procesorul ales pentru implementare. Pentru familia Texas Instruments structura este bazată pe conceptul DSP BIOS [1], generalizat în aplicațiile cu DSP din generația 6000.

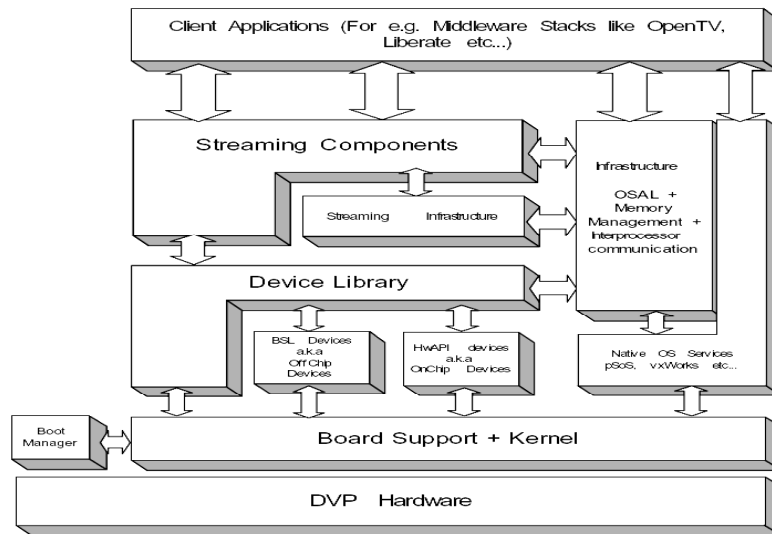


Figura 7. Arhitectura TSSA ([2])

Philips Nexperia este bazat pe arhitectura software TSSA (Trimedia Streaming Software Architecture [2]). In această arhitectură aplicația propriu-zisă este doar o parte mică din programul sistem. In schimb, suportul oferit de arhitectura TSSA permite accesul la biblioteci optimizate de prelucrare de imagini, și la o modalitate transparentă a schimbului de date din sistem. O imagine tipică a structurii unei aplicații pe sisteme Nexperia este dată în figura 7.

In principal blocurile software necesare a echipa un sistem de recunoaștere a fețelor (de tip server) sunt următoarele:

- Nucleu sistem de operare
- Drive de achiziție
- Drive pentru normalizarea imaginii
- Algoritm de recunoaștere a feței
- Clasificare
- Gestiune alarme
- Drive de comunicație

Structura modului client

Programul care este utilizat în sistemul de dispecerizare al rețelei de camere de recunoaștere, are componente care comandă câteva funcții esențiale:

- Poziționare, comandă zoom optic
- Gestiunea alarmelor
- Descărcarea imaginilor de referință
- Descărcare imagini pentru arhivare

PERFORMANȚE ȘI EVALUARE

Aplicația prezentată în [6] a fost testată pe o imagine de test de 640x480 pixeli care conține o singură față. Baza de date este compusă din 10 fețe. Datorită dimensiunilor mici ale bazei de date căutarea se face în memoria sistemului. Variabilitatea timpului de calcul în funcție de structura fișierului a necesitat medierea pe mai multe cicluri de calcul. Rezultatele sunt

prezentate în tabelul 2. O scurtă privire ne arată că detecția feței consumă cel mai mare timp de calcul, iar blocurile de clasificare cel mai mare volum de memorie.

O creștere în dimensiunea bazei de date va crește necesitățile de procesor și de memorie pentru blocurile de clasificare. O examinare a rezultatelor pentru o imagine statică arată că sunt necesare mai puțin de 4 secunde pentru a recunoaște o față.

	Număr de cicluri (x10 ⁶)	Timp de calcul (CPU la 500MHz)	Consum memorie
Detecție față	1161	2.32 s	~ 392 KB
Localizare ochi	585	1.17 s	~ 436 KB
Normalizarea feței	56	0.11 s	~ 32 KB
Clasificarea feței (metoda Eigenfaces)	18	0.04 s	~ 1055 KB
Clasificarea feței (metoda Segmentare liniară)	22	0.05 s	~ 2064 KB

Tabelul 2- Rezultate ale implementării algoritmilor de recunoaștere a fețelor

Majoritatea timpului este consumat pentru detecția feței și localizarea ochilor. Algoritmi mai rapizi pentru aceste etape pot îmbunătăți timpul de prelucrare în mod semnificativ. Deși nu s-a făcut implementarea algoritmului bazat pe recunoașterea din secvențe video, estimarea realizată pentru imagini statice arată că este posibilă o implementare în timp real pe TMS320C64x. Într-un astfel de sistem care urmărește un personaj căutarea și localizarea feței se fac doar la începutul secvenței, normalizarea și clasificarea durează mai puțin de 1 secundă. Concluzia este că procesoarele investigate pot fi utilizate în implementarea unui sistem de recunoaștere în timp real. Problema poate fi mutată la serverul central, care poate primi doar imaginea localizată și normalizată, procesul de recunoaștere fiind realizat de către server. Dimensiunea mare a blocurilor de clasificare și a memoriei cerute de algoritmi poate face atractivă o astfel de soluție.

În ceea ce privește eficiența algoritmilor se observă că metoda „Eigenfaces” este mai rapidă, consumă mai puțină memorie, dar asigură o rată de recunoașteri mai mică decât metoda subspațiilor segmentate liniar. Implementarea trebuie să țină cont de aceste caracteristici și să facă, dacă este cazul, compromisuri.

CONCLUZII

Problema recunoașterii fețelor este una din cele mai complexe probleme din cadrul recunoașterii imaginilor. Ea implică cunoașterea anatomiei feței, și necesită algoritmi de preprocesare și clasificare extrem de complecși.

Implementarea unor sisteme autonome de recunoaștere a fețelor este încă la începuturile sale. Extinderea rapidă este împiedicată de dificultățile care există încă în algoritmi de prelucrare și rata de recunoaștere relativ redusă. În aceste cazuri transmisia imaginilor în dubiu pentru analiza „off line” de către operatorul uman simultan cu recunoașterea automată, poate fi o opțiune pentru evitarea recunoașterilor false.

Puterea de prelucrare a unor procesoare DSP de înaltă performanță, procesoarele media, permit implementarea eficientă din punct de vedere a costului, dimensiunilor și consumului de putere. Perfecționarea algoritmilor de prelucrare va permite ca aceste sisteme să devină o alternativă la binecunoscutele camere de supraveghere actuale.

Bibliografie

- [1] * * *, "TMS320DM642 Video/Imaging Fixed-Point Digital Signal Processor, *Data Manual*, Texas Instruments", 2002
- [2] * * *, "PNX1300 Series Media Processors", *Preliminary Specification*, Philips Semiconductor, 2002
- [3] Д.Брилюк, В. Старовойтов, „Методы распознавания человека по изображению лица. Достоинства и недостатки, сравнение”, *website: <http://daily.sec.ru>*, 2002
- [4] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1991, pp. 586-591.
- [5] A. U. Batur and M. H. Hayes. "Linear Subspaces for Illumination-Robust Face Recognition", in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2001, pp.296-301.
- [6] Aziz U. B. and B. E. Flinchbaugh, „Performance Analysis of Face Recognition Algorithms on TMS320C64x”, *SPRA874, Texas Instruments*, dec. 2002
- [7] R. Arsinte, "Face recognition module" , *Technical report* , Microtech, 2003