

Univestitatea POLITEHNICA Bucuresti

Facultatea de Automatica si Calculatoare
Specializarea Calculatoare si Tehnologia Informatiei



Voting-Based Image Binarization
Software Design Document
Version 1.1
(Managementul Proiectelor Software)

Indrumator:

Saca Octavian

Componenta echipei:

Costea Cristina 341 C5
State Andra 341 C5
Barahtian Oana 341 C5

Bucuresti
2013

Cuprins

1. Rolul documentului
2. Continutul documentului
3. Principalele rolurile in echipa
 - 3.1. Dezvoltator (Costea Cristina)
 - 3.2. Tester (State Andra)
 - 3.3. Project Manager (Barahian Oana)
4. Despre proiect
 - 4.1. Scop
 - 4.2. Context
5. Abordarea proiectului
 - 5.1. Modelul datelor
 - 5.2. Elemente de testare
 - 5.2.1. Componente critice
 - 5.2.2. Tipuri de testare
 - 5.2.3. Alternative
6. Algoritm
7. Responsabilitati

1. Rolul Documentului

Documentul actual vine in completarea codului sursa uploadat pe GitHub, explicand abordarea si constructia proiectului, coluziunea in echipa, punctand si momentele cheie ale proiectului cat si concluziile.

Acest document ii ajuta pe colegii din celelalte mini-echipe sa inteleaga abordarea cat si structura codului sursa.

2. Continutul documentului

Documentul este structurat in 4 sectiuni esentiale:

1. Organizarea echipei
2. Detalii de implementare
3. Modelul datelor
4. Modelul architectural al datelor si elemente de testare

3. Rolurile in echipa

Fiecare mini-echipa are in component un dezvoltator (Dev), un tester si un project manager (PM). Structurarea echipei aduce un plus de organizare si de usurinta a impartirii sarcinilor, acestea facandu-se pe puncte cheie din proiect.

Fiecare veriga are responsabilitati conturate mai mult mai putin concret, anumite etape fiind comune – spre exemplu partea e research, fiind essential ca fiecare membru al echipei sa inteleaga cerintele si specificatiile. De asemenea, un alt punct important este deciderea abordarii de lucru. In cele ce urmeaza voi detalia responsabilitatile fiecareia dintre noi.

3.1. Dezvoltatorul - Costea Cristina

Cristina a avut ca responsabilitate principala implementarea software a cerintei. Intrucat partea de implementare a depins in mare parte de ea, dupa cateva sesiuni de brainstorming, decizia finala asupra algoritmului folosit cat si modul de abordare i-au apartinut.

3.2. Tester – State Andra

Rolul testerului este de a verifica functionalitatea aplicatiei si identificarea neregulilor.

3.3. Project Manager – Barahtian Oana

Desi nu este implicat in mod direct in procesul de implementare, PM-ul este responsabil de succesul/esecul proiectului. Ca responsabilitate principala ar putea fi mentionata planificarea si monitorizarea evolutiei proiectului.

4. Despre proiect

4.1. Scop

Scopul aplicatiei rezultate este digitalizarea unor documente prin supunerea lor unor algoritmi complexi de binarizare.

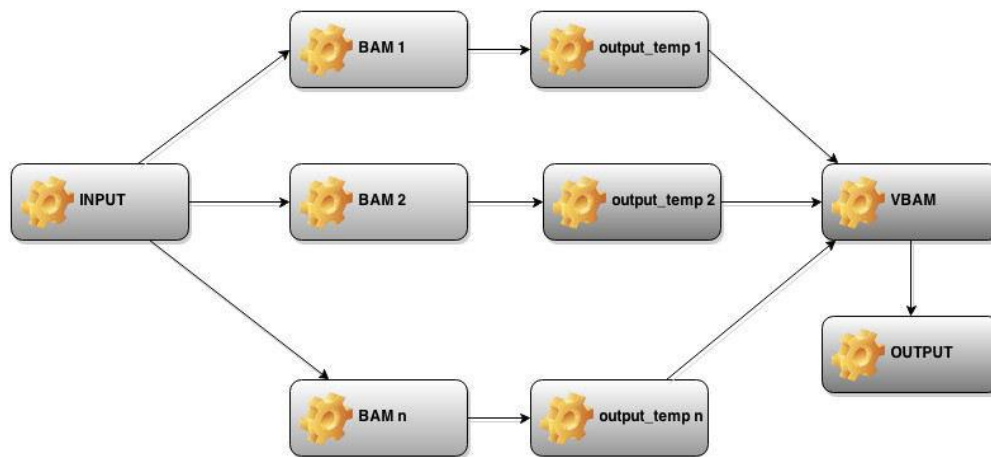
Documentele candidate sunt din seria celor din deceniile trecute, ce se gasesc doar in forma fizica, cu o lizibilitate scazuta. Acestea pot fi manuscrise sau tiparituri ce contin informatii importante ce s-ar dori sa fie procesate in mod automat.

4.2. Context

Proiectul are drept scop dezvoltarea unui sistem IBS (de binarizare a imaginilor) ce include doua module:

- Modulul de binarizare care transfera imaginea in una binara
- Modulul de votare care compune o singura imagine binara folosind mai multe rezultate ale mai multor module de binarizare.

4.3. Diagrama de arhitectura



5. Abordarea proiectului

In acest capitol vor fi prezentate succint structurile de date folosite, formatul fisierelor cat si elementele de testare.

5.1. Modelul datelor

Proiectul are ca date de input, fisierul ce urmeaza a fi binarizat. Acest fisier este retinut intr-o structura globala. Incarcarea se realizeaza o singura data. Imaginile sunt stocate in matrici alocate dinamic.

Modulul de care ne-am ocupat fiind cel de binarizare, outputul aplicatiei consta in doua imagini: prima reprezinta solutia dupa aplicarea algoritmului, iar cea de-a doua reprezinta factorul de incredere pentru corectitudinea imaginii output.

5.2. Formatul fisierelor

Imaginile vor avea format de tip .PGM, iar un fisier de tipul .PGM va contine:

- Un numar magic pentru identificare fisierului (Ex: P2/P4);
- Lungimea si latimea reprezentate de numere N (numarul de linii a matricei de pixeli) si M (numarul de coloane);
- Matricea de pixeli propriu-zisa.

Ex:

P2/P4

```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 3 3 3 3 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 15 0
0 3 3 3 0 0 0 7 7 7 0 0 0 11 11 11 0 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 0 0
0 3 0 0 0 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

5.3. Elemente de testare

Pentru verificarea corectitudinii aplicatiei se introduc parametrii de testare gresit, imagini corupte.

5.3.1. Componente critice

Pentru a se verifica corectitudinea aplicatiei se vor introduce parametrii de intrare gresit, se vor introduce fisiere de un alt format, altul decat cel suportat de aplicatie, imagini corupte, numar gresit de BAM-uri. Pentru testare se va realiza de asemenea un script care sa preia automat datele din fisierele de intrare, sa le prelucraze și sa afișeze imaginile corecte intr-un fisier de output. Fiecare componenta va fi testata individual.

5.3.2. Tipuri de testare

Testarea se poate realiza ca:

1. Testare White-Box ce presupune ca testerul care acces in sursele programului. De multe ori, testarea prin aceasta metoda implica scrierea de cod sau cel puțin, urmarirea celui existent. Se testeaza fiecare metoda in parte inițial, cât si interoperarea metodelor.
2. Testare Black-Box ce presupune testarea aplicatiei la nivelul user-ului, plecand de la premisa ca nu se cunoaste modul in care functioneaza. Pentru acest tip de testare, testerul trebuie sa cunoasca rezultatele ce se așteapta din partea aplicației.

5.3.3. Alternative

Modulul VBAM poate executa in paralel evaluarea fiecarui modul BAM prin pornirea mai multor procese.

6. Algoritm

Procesul de research a fost unul interesant. In urma acestui process, candidatii cei mai viabili erau Nina, Niblack Sauvola si Wolf. In urma testarii celor 3 algoritmi, in functie de acuratetea output-ului, cat si ca formula matematica aplicata, Wolf a fost desemnat castigator. Wolf are avantajul calcularii unui prag local, ceea ce duce la o imagine mai coerenta.

7. Responsabilitati

Nr. crt	Activitate	Numar ore estimate	Numar ore allocate	Responsabil
1.	Research - Dev	10	12	Costea Cristina
2.	Research - Tester	10	8	State Andra
3.	Research - PM	10	6	Barahitian Oana
4.	Alegerea unui algoritm	4	5	Costea Cristina Barahitian Oana

5.	Familiarizarea cu framework-ul	2	4	Costea Cristina
6.	Implementarea efectiva	8	12	Costea Cristina
7.	Crearea scenariilor de test	4	3	State Andra
8.	Testare efectiva	6	8	State Andra Barahtian Oana
9.	Optimizare si rezolvarea bug-urilor	5	4	Costea Cristina
10.	Creare SDD v1.0	3	5	State Andra
11.	Completare raport saptamanal	2	2	Barahtian Oana
12.	Creare README (SDD v1.1)	2	2	Barahtian Oana
13.	Code review	2	3	State Andra Barahtian Oana
14.	Creare prezentare finala			Barahtian Oana