



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI



Instrumente Structurale  
2007-2013



# Platformă de e-learning și curriculum e-content pentru învățământul superior tehnic

**AEACD**

## **20. Analiza de layout**

# Layout

- Detectia de layout a unei pagini este importanta pentru clasificarea elementelor existente
- Elementele de layout:
  - Imagini
  - Paragrafe
  - Caractere special (numarul paginii)

# Paragrafe si Imagini

- Paragrafele si imaginile sunt detectate folosind elementele mentionate in capitolele anterioare: separatori, linii, spatii albe
- Se folosesc metodele deja prezentate peste care se aplica o noua metoda de clasificare a paragrafelor si imaginiiilor
- Metode noi folosesc Triangularizarea Delaunay sau distantele existente in imagine

# Distante

- Distanțele între entități pot ajuta la detectia de layout
- Pagina va fi formata din zone (cluster)
- Doua tipuri de distante: euclidiană și ierarhică
- Algoritmul folosește ambele tipuri
- Distanța ierarhică reprezintă distanța maximă din interiorul unui cluster

# Distante

- Algoritmul incepe prin calcularea tuturor distantelor euclidiene intre entitati si sortarea lor in ordine crescatoare
- Se porneste de la cea mai mica distanta si se unesc intr-un cluster entitatile conectate de acea distanta
- Se repeta procesul pentru urmatoarea cea mai mica distanta
- Daca nu se creaza un punct de oprire algoritmul se va opri in momentul in care toata pagina va fi un cluster

# Distante - Selectia rezultatelor

- Pentru a determina punctul de oprire se calculeaza la fiecare pas ariile dreptunghiurilor de incadrare ale clusterelor
- Trei tipuri de arie: totala, dreptunghiurilor suprapuse si dreptunghiurilor nesuprapuse
- Se realizeaza histograma valoriiilor celor trei si punctul in care aria dreptunghiurilor nesuprapuse este maxim cand aria dreptunghiurilor suprapuse este 0, minima sau foarte mica este punctul de oprire
- Rezultatul final vor fi clusterelor ce reprezinta zonele de layout din imagine

# Triangularizarea Delaunay

- Scopul metodei este de a uni fiecare trei puncte de pe conturul unei entitatii astfel incat orice triunghi se formeaza sa nu intersecteze alt punct



# Detectia de separatori

- Primul pas este detectia de separatori
- Se aplica triangularizarea Delaunay in imagine
- Se pastreaza doar triunghiurile ce unesc fix doua entitati
- Pentru fiecare entitate se calculeaza raportul dintre numarul de puncte Delaunay aflate pe entitate si numarul de puncte Delaunay aflate pe celelate entitati conectate cu cea curenta
- Raportul va fi mai mare, valoare mai mare ca 1, pentru separatori – imagini sau linii





I.





Von der

**E i n t h e i l u n g**

des deutschen Reichs in Franken  
und Sachsen.

---





Diferite tipuri de entitati si valorile raportului pentru fiecare folosind algoritmul pentru detectia de separatori

Entitate	Raport
	2.7328
	3.2832
	0.7610
	0.9662





# Detectia de separatori

- Aceasta metoda poate fi imbunatatita in functie de rezultatul dorit
- Raportul se inmulteste cu aria separatorului pentru a detecta in special imaginile
- Raportul se inmulteste cu gradul de umplere (fill ratio) al separatorului pentru a detecta liniile

Diferite tipuri de entitati si valorile raportului inmultit cu aria entitatii

Entitate	Raport
	2102128
	309269
	6810
	5293

Diferite tipuri de entitati si valorile raportului inmultit cu gradul de umplere

Entitate	Raport
	0.884348
	1.680385
	0.6445
	0.6916

# Detectia de separatori

- Detectia separatoriilor spatii albe se face folosind o metoda deja mentionata, dar folosita anterior cu un alt scop
- Se realizeaza un resampling pe imagine initiala, iar apoi se coloreaza cu negru tot ceea ce nu este alb clar
- In final se vor obtine separatorii spatii albe



### THE STORY OF THE THREE PIGS

Once there was a mother pig  
who had three little pigs.

Their names were Whitey, Brownie,  
and Blacky.

The three little pigs grew and grew.

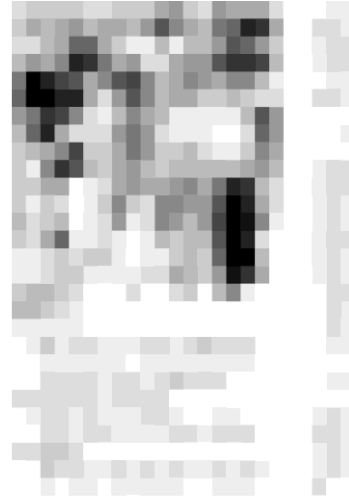
At last they were  
to go away from home  
So their mother  
and sent them away

WENT

Whitey met a  
wolf with a bundle of  
sticks. He said, "Please  
will you give me  
to make me a house."  
"Yes, you may."  
"Thank you."  
Then he made

THE WOLF

Soon the wolf  
and rapped at the  
house. He said, "Little  
pig, let me in."



# Paragrafele si imaginile - Combinarea metodelor

- Avand detectate atat spatiile albe cat si separatorii, revenind la triangularizarea Delaunay initiala, toate muchiile ce unesc separatori si ce trec peste spatiile albe detectate sunt eliminate
- In final paragrafele vor fi constituite din entitatiile legate prin muchiile triangularizarii Delaunay





### THE STORY OF THE THREE PIGS

Once there was a mother pig  
who had three little pigs.

Their names were Whitey, Brownie,  
and Blacky.

The three little pigs grew and grew.

At last they were  
to go away from home.  
So their mother  
and sent them away.

Whitey

Whitey met a  
with a bundle of  
He said, "Pie  
will you give me

to make me a hot  
"Yes, you may  
"Thank you.

Then he made

THE WOLF

Soon the wolf  
and rapped at the  
He said, "Li

let us