

---

Súlyozott automaták alkalmazása  
képek reprezentációjára,  
képtranszformációra és beszédfelismerésre

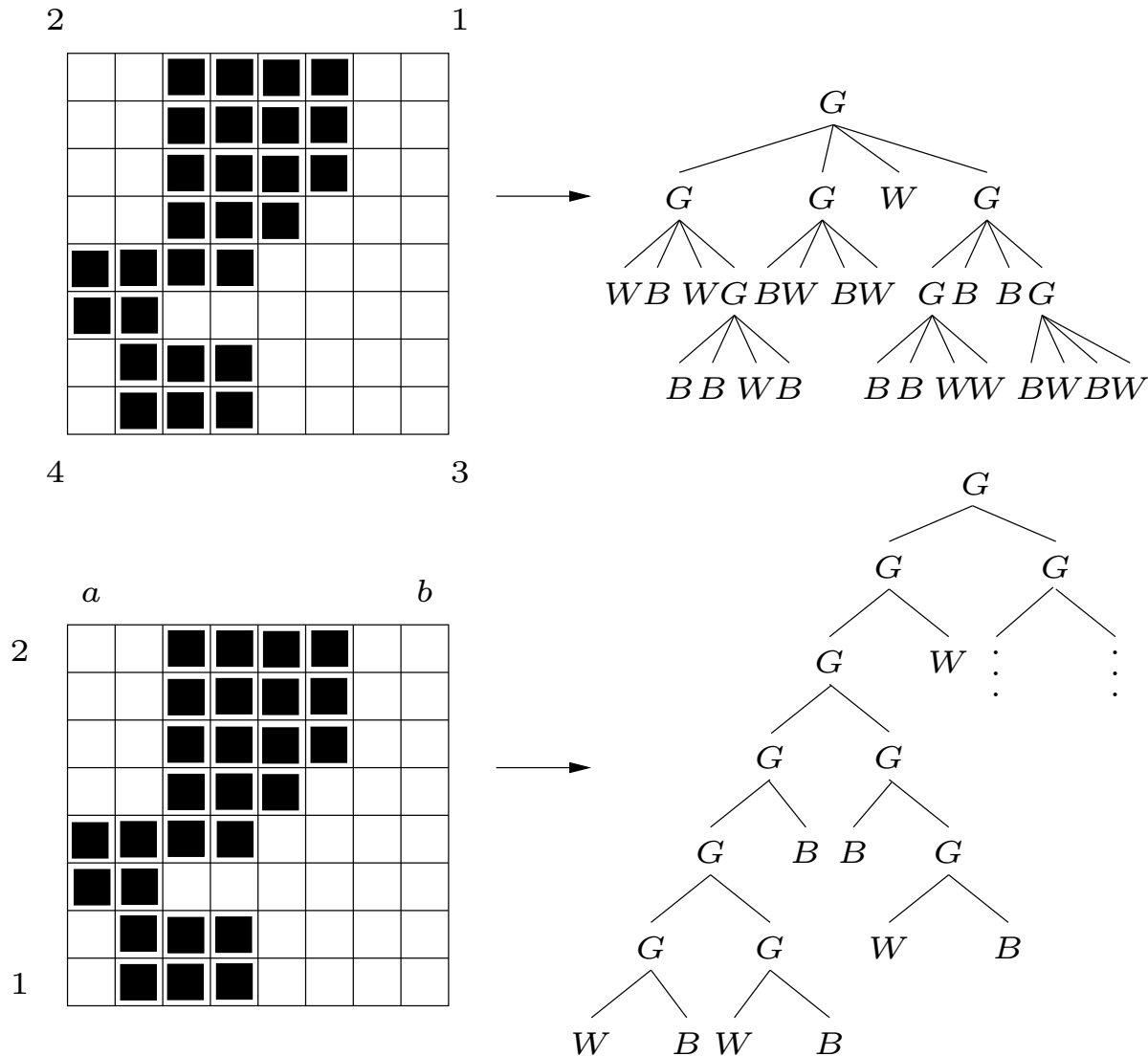
Gazdag Zsolt

Szegedi Tudományegyetem  
Számítástudomány Alapjai Tanszék

---

- Képek kódolása és dekódolása  
(az előző előadás folytatása)
- Képek transzformációja
- Beszédfelismerés

# További képfelosztási módszerek



Egy kép és az őt kódoló négyágú illetve bináris fa

Amikor a vízszintes és függőleges felosztás nem váltakozva történik.



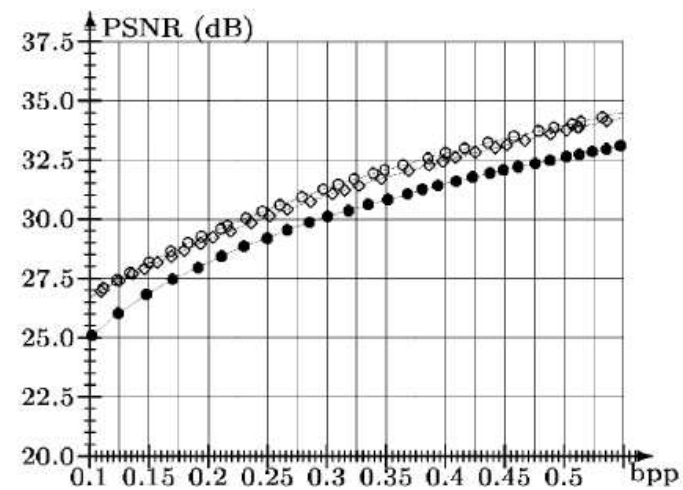
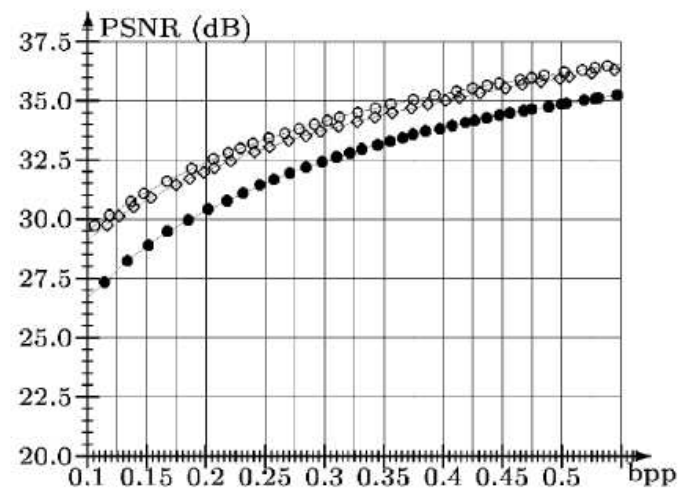
bináris felosztás



light HV felosztás

Az eredeti HV felosztás nem gazdagságos (több információt kell tárolni, mint amit nyerünk a módszerrel)

# A „light” HV felosztás [KMT04]



○ : light HV, ◇ : bintree, ● : IJPEG

# Képek dekódolása

---

„Direkt” dekódolás:  $2^n \times 2^n$  kép dekódolása  $O(|Q|(|\Sigma||Q|)^n)$  idejű

Egy gyorsabb algoritmus ( $O(|Q|^2|\Sigma|^n)$  idejű):

**Input:** Egy  $\mathcal{A}$  automata és egy  $n$  szám

**Output:** Az  $\mathcal{A}$  által reprezentált kép  $n \times n$  felbontásban

**For** ( $i = 1; i \leq |Q|; i++$ )  $f_i(\varepsilon) = F_i;$

**For** ( $l = 1; l \leq n; l++$ )

**For** ( $i = 1; i \leq |Q|; i++$ )

**For** ( $a \in \Sigma$ )

**For** ( $w \in \Sigma^{l-1}$ )

$$f_i(aw) = \sum_{j=1}^{|Q|} W_a(i, j) f_j(w);$$

**For** ( $w \in \Sigma^n$ )

$$f_{\mathcal{A}}(w) = \sum_{i=1}^{|Q|} I_i f_i(w);$$

Legyen  $f_{\mathcal{A}}(\Sigma^n)$  a dekódolt kép



128 × 128 méretű Lena nagyítása  
512 × 512 méretre sv-automatával



128 × 128 méretű Lena 512 × 512  
méretre nagyítva

- Az sv-automaták alkalmasak képek tömörítésére
- A képek kódolása nagyon időigényes
- Alkalmasak képek és videók tárolására pl. CD-ROM-on
- Nem alkalmasak szimmetrikus képátvitelre (pl. videokonferencia)



*Súlyozott véges transzformátor (sv-transzformátor):*

$$\mathcal{M} = (Q, \Sigma_1, \Sigma_2, W, I, F)$$

- $Q$ : véges állapothalmaz
- $\Sigma_1, \Sigma_2$ : input és output ábécék
- $W : Q \times (\Sigma_1 \cup \{\varepsilon\}) \times (\Sigma_2 \cup \{\varepsilon\}) \times Q \rightarrow \mathbb{R}$ , a súlyfüggvény  $((W_{a,b})_{i,j} = W(i, a, b, j))$
- $I, F : Q \rightarrow \mathbb{R}$ , a kezdő illetve befejező hozzárendelések

Az  $\mathcal{M} = (Q, \Sigma_1, \Sigma_2, W, I, F)$  által indukált

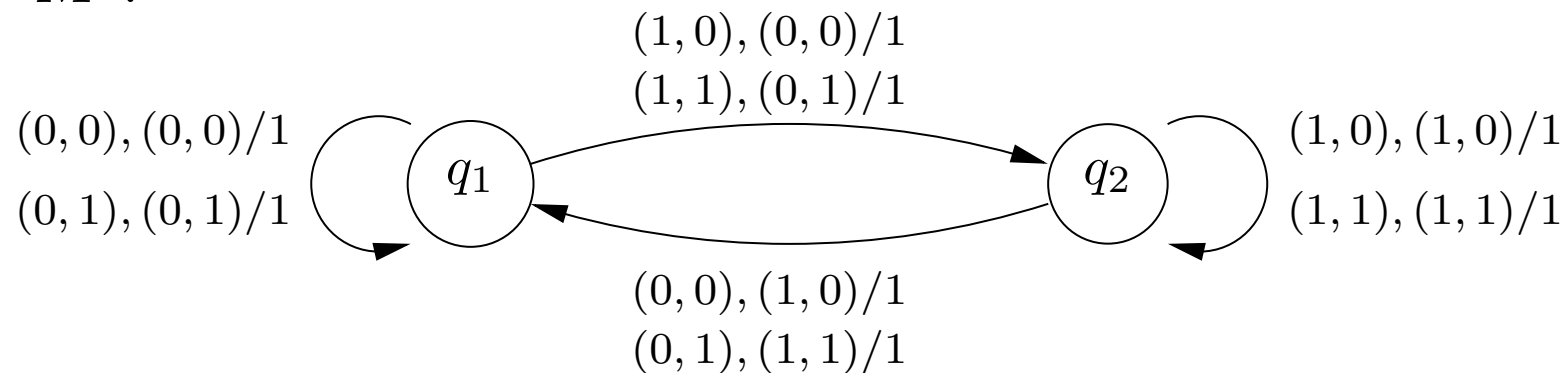
$f_{\mathcal{M}} : \Sigma_1^* \times \Sigma_2^* \rightarrow \mathbb{R}$  *súlyozott reláció* ( $\varepsilon$ -mentes esetben):

$$f_{\mathcal{M}}(a_1 a_2 \dots a_k, b_1 b_2 \dots b_k) = I^t W_{a_1, b_1} W_{a_2, b_2} \dots W_{a_k, b_k} F$$

Ha  $\rho : \Sigma_1^* \times \Sigma_2^* \rightarrow \mathbb{R}$  egy súlyozott reláció és  $f : \Sigma_1^* \rightarrow \mathbb{R}$  egy többszörös felbontású függvény akkor  $g = \rho(f) : \Sigma_2^* \rightarrow \mathbb{R}$  :

$$g(v) = \sum_{u \in \Sigma_1^*} f(u) \rho(u, v)$$

$M$  :



$$I = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, F = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$f_{\mathcal{M}}((1, 1)(1, 0), (0, 1)(1, 0)) = [0.5, 0.5] \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.5$$

1	$\frac{5}{4}$	$\frac{3}{2}$	2
$\frac{3}{4}$	1	$\frac{5}{4}$	$\frac{3}{2}$
$\frac{1}{2}$	$\frac{3}{4}$	1	$\frac{5}{4}$
$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1

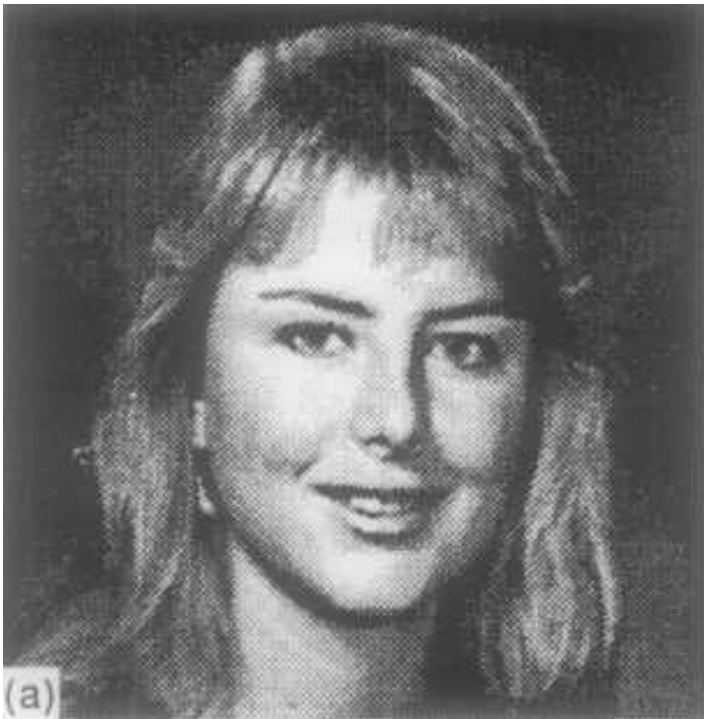
$$z = x + y$$

Ha  $f$  a  $z = x + y$  függvényhez konvergáló többszörös felbontású függvény és  $g = f_M(f)$ , akkor

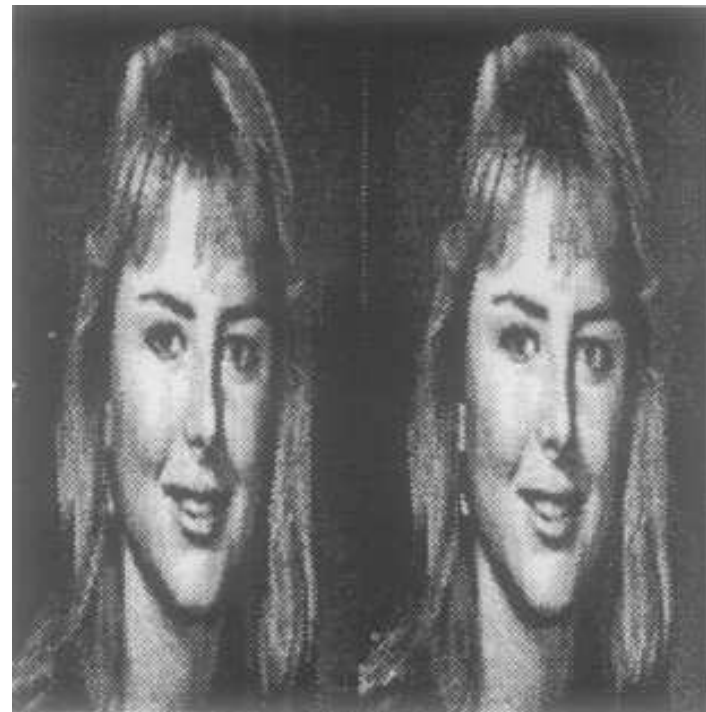
$$\begin{aligned}
 g((0,0)(0,0)) &= \sum_{u \in \Sigma^*} f(u) \rho(u, (0,0)(0,0)) = \frac{3}{8} \\
 g((0,0)(1,0)) &= \sum_{u \in \Sigma^*} f(u) \rho(u, (0,0)(1,0)) = \frac{7}{8} \\
 g((1,0)(0,0)) &= \sum_{u \in \Sigma^*} f(u) \rho(u, (1,0)(0,0)) = \frac{3}{8} \\
 g((1,0)(1,0)) &= \sum_{u \in \Sigma^*} f(u) \rho(u, (1,0)(1,0)) = \frac{7}{8}
 \end{aligned}$$

$\frac{3}{8}$	$\frac{7}{8}$	$\frac{3}{8}$	$\frac{7}{8}$

$$(x, y) \rightarrow \left\{ \left( \frac{x}{2}, y \right), \left( \frac{x+1}{2}, y \right) \right\}$$



*Carol* többszörös felbontású kép



$f_M(\textit{Carol})$

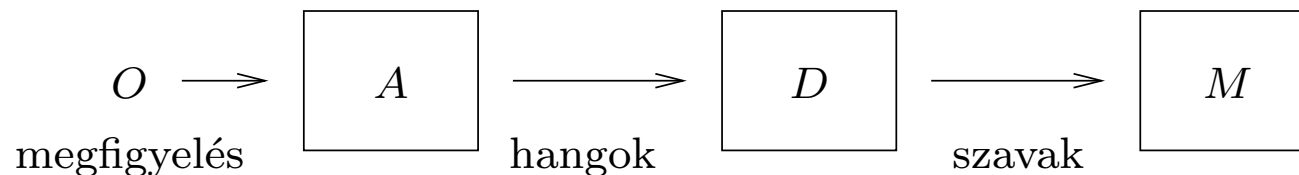
# Beszéd felismerés súlyozott transzformátorokkal

---

**A beszéd felismerés feladata:** Adott  $u$  akusztikus események sorozata. Keressük meg azt a  $w$  szót amelyre a legnagyobb a valószínűsége annak, hogy  $w$  az  $u$  sorozatot indukálja.

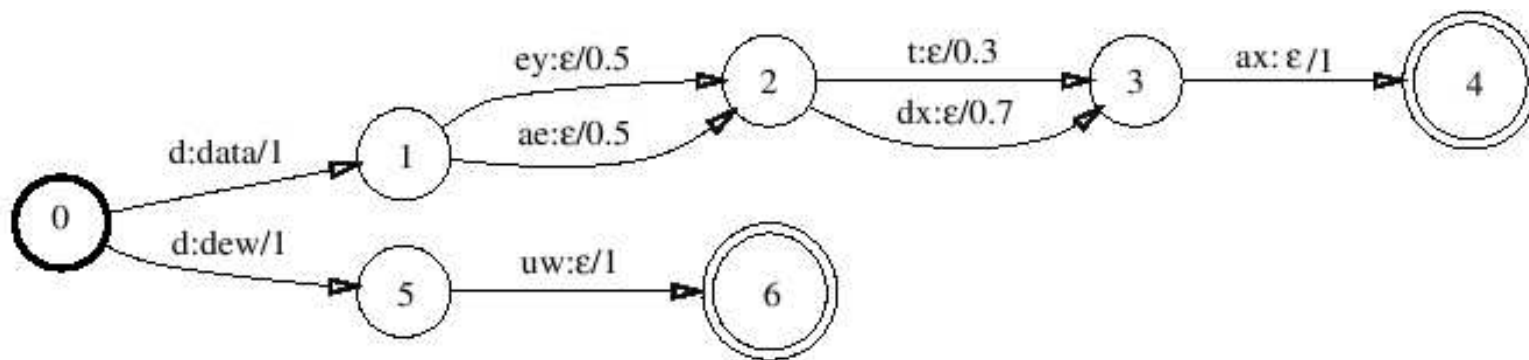
**Lehetséges megoldás:** Valamilyen statisztikai modell használata

Ilyen modell lehet súlyozott transzformátorok kompozíciója:



- $O$ : Súlyozott automata ami a beszédforrást ún. akusztikus események sorozataként ábrázolja
- $A$ : Súlyozott transzformátor, akusztikus megfigyelési sorozatokat képez hangsorozatokba
- $D$ : Súlyozott transzformátor, hangsorozatokat képez szavak sorozatába
- $M$ : Nyelvi modell, szavak (vagy betűk) sorozatainak valószínűségét adja

Egyszerűsített példa az  $M$  transzformátor egy résztranszformátorára [MPR02]:



A kezdőállapotból a végállapotig haladva megkapjuk a *data* szó egy kiejtésének a valószínűségét

A  $\mathbb{K}$  félgűrű feletti  $\rho_1 : \Sigma_1^* \times \Sigma_2^* \rightarrow \mathbb{K}$  és  $\rho_2 : \Sigma_2^* \times \Sigma_3^* \rightarrow \mathbb{K}$  súlyozott transzformációk *kompozíciója*:

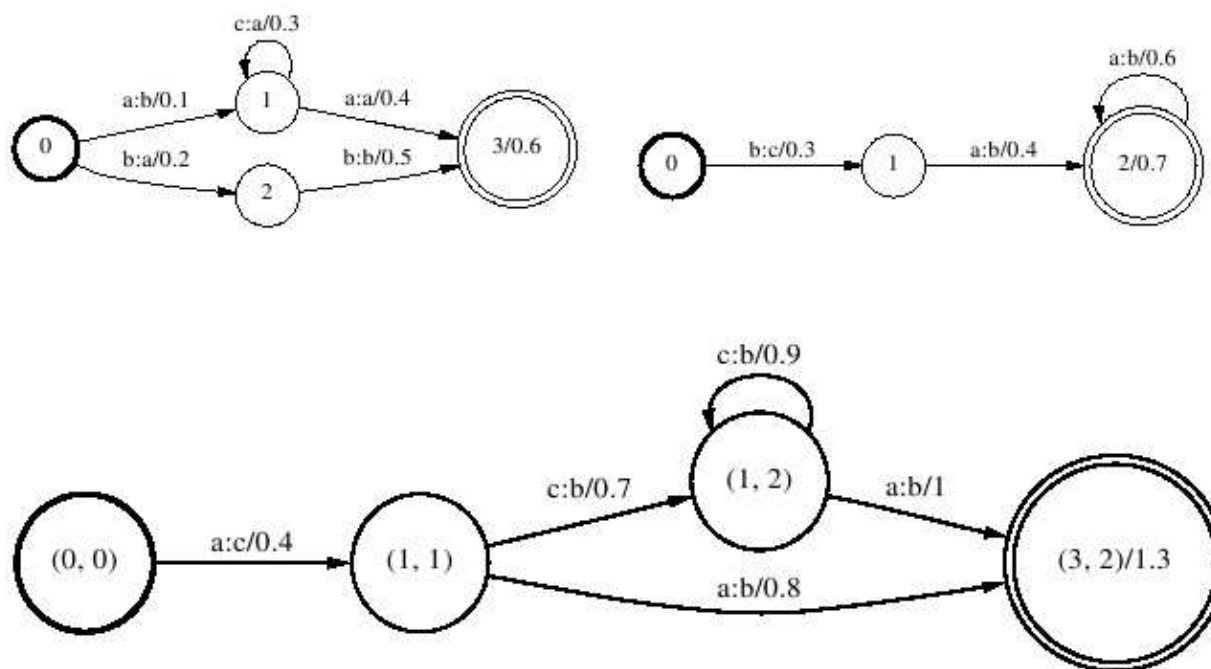
$$(\rho_1 \circ \rho_2)(u, w) = \sum_{v \in \Sigma_2^*} \rho_1(u, v) \rho_2(v, w)$$

A súlyozott transzformátorok kompozícióját úgy kell megadni, hogy az általuk indukált transzformációk kompozíciója a kapott súlyozott transzformátor által indukált transzformáció legyen.



# Súlyozott transzformációk és transzformátorok kompozíciója

**Példa:**  $Trop = (\mathbb{R}_+ \cup \{\infty\}, \min, +, \infty, 0)$  félgűrű feletti súlyozott transzformátorok kompozíciója [MPR02]:



A súlyozott transzformátorokkal történő beszéd felismerés előnye: A gyakorlatban általában minden transzformátor kompozíció után hatékonyan elvégezhetőek az alábbi egyszerűsítések (lásd pl. [Moh97]):

- A kapott transzformátor determinizálása
- A kapott determinisztikus transzformátor minimalizálása

Igy a beszéd felismerést végző transzformátor (az  $O$ ,  $A$ ,  $D$  és  $M$  transzformátorok kompozíciója) hatékony marad (idő és tár tekintetében).

**K. Culik II és J. Kari.** Image compression using weighted finite automata. In *MFCS*, 392–402, 1993.

**K. Culik II és J. Kari.** Finite State Transformation of Images. *Comput. & Graphics*, 20, 125–135, 1996.

**F. Katritzke, W. Merzenich, M. Thomas.** Enhancements of partitioning techniques for image compression using weighted finite automata. *Theoretical Computer Science*, 313, 133–144, 2004.

**F. Katritzke.** Refinements of data compression using weighted finite automata. Ph.D. Thesis, University of Siegen, 2001.

**M. Mohri.** Finite-State Transducers in Language and Speech Processing. *Computational Linguistics*, 23, 269–311, 1997.

**M. Mohri, F.C.N. Pereira és M. Riley.** Weighted Finite-State Transducers in Speech Recognition. *Computer Speech & Language*, 16, 69–88, 2002.

**F.C.N. Pereira és M.D. Riley.** Speech Recognition by Composition of Weighted Finite Automata. 1996.