LMD konf. darbai, 336–340
1997, Vilnius

Apie baigtinių automatų teorijos taikymą apdorojant dvimačius vaizdus

J. Valantinas (KTU)

Viena iš svarbesnių realaus pasaulio vaizdų savybių yra ta, kad jiems galima sudaryti bet kurio jų detalizacijos lygio matematinį modelį. Imkime vaizdą $V \in R$, R – realaus pasaulio vaizdų (su „pilka" šviesos intensyvumo skale) erdvė. Šio vaizdo s-detalizacijos lygio ($s = 0, 1, 2, \ldots$) matematinį modelį V_s sutapatinsime su dvimačiu duomenų masyvu $[V_s(\omega)]$, kurio matmenys $(2^s \times 2^s)$; be to, priimsime, jog kiekvieną masyvo elementą (pikselį) $V_s(\omega)$ atitinka vaizdo V fragmentas (blokelis), „lokalizuojamas" adresu ω,

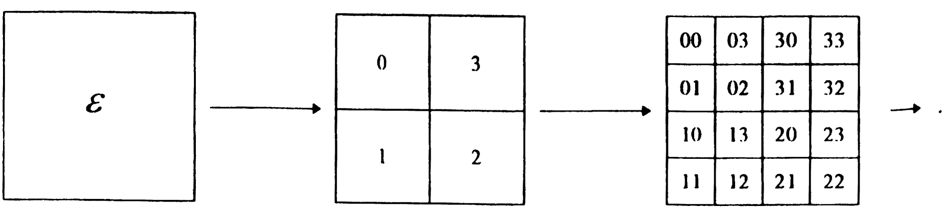
$$\omega = \{\omega_1\omega_2\cdots\omega_s \mid \omega_i \in \{0, 1, 2, 3\};\ i = 1, 2, \ldots, s\}$$

(adreso ω ilgis $|\omega| = s$), bei charakterizuojamas fiksuota šviesos intensyvumo skaitine reikšme iš aibės $\{0, 1, 2, \ldots, 255\}$. Keli pirmieji vaizdo V žemiausiojo detalizacijos lygio ($s = 0, 1, 2$) matematiniai modeliai (V_0, V_1, V_2) bei juos sudarančių blokelių adresavimo tvarka parodyta pav. 1.

Sąryšis tarp bet kurių dviejų vaizdo V skirtingo detalizacijos lygio matematinių modelių, V_s ir V_t, $0 \leqslant s < t$, išreiškiamas tokia lygybe:

$$\forall\omega\big(|\omega| = s\big)\colon\ V_s(\omega) = \text{apv.}\left(\frac{1}{4^{t-s}} \sum_{\tau\in\tau(\omega)} V_t(\tau)\right);$$

čia: $\tau(\omega) = \{\omega_1\omega_2\cdots\omega_s\sigma_1\sigma_2\cdots\sigma_{t-s} \mid \sigma_j \in \{0, 1, 2, 3\};\ j = 1, 2, \ldots, t-s\}$, ir apv.$(x)$ reiškia skaičiaus $x \in \mathbf{R}$ apvalinimą iki artimiausiojo sveikojo skaičiaus.



0 - detalizacijos lygis
(V_0- masyvas (1 x 1));
ε - modelio V_0, kaip atskiro elemento, adresas; $|\varepsilon| = 0$

1 - detalizacijos lygis
(V_1 - masyvas (2x2))

2 - detalizacijos lygis
(V_2- masyvas (4x4))

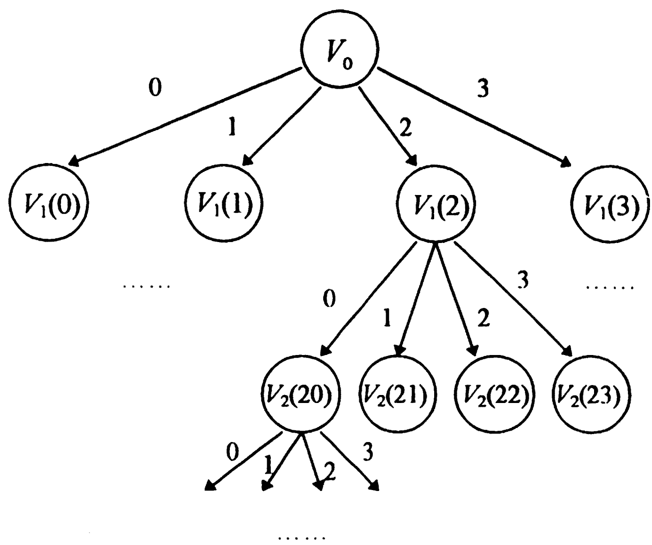
Pav. 1.

Pastebėsime, jog vaizdo $V \in R$ bet kurio detalizacijos lygio matematinį modelį galima taipogi aprašyti orientuotu grafu-medžiu, kurio kiekvienai viršūnei priskiriama konkreti skaitinė intensyvumo reikšmė iš aibės $\{0, 1, 2, \ldots, 255\}$. Pradinę viršūnę atitinka vaizdo modelis V_0. Iš bet kurios grafo viršūnės išeina keturi grafo lankai, pažymėti simboliais 0, 1, 2 arba 3, pav. 2.

Iš grafiko (pav. 2) matyti, kad kelias iš bet kurios grafo-medžio viršūnės, esančios s-detalizacijos lygyje, į bet kurią kitą viršūnę, esančią t-detalizacijos lygyje ($t > s \geqslant 0$), vienareikšmiškai nusakomas $(t - s)$ ilgio adresu σ, $\sigma = \{\sigma_1\sigma_2 \cdots \sigma_{t-s} \mid \sigma_j \in \{0, 1, 2, 3\}$; $j = 1, 2, \ldots, t - s$.

Aišku, jog, realiai imant, vaizdo V matematinio modelio V_s detalizacijos lygis s negali būti begalinis, t. y. praktiniuose taikymuose šis dydis yra aprėžtas iš viršaus, būtent: $s \leqslant n$, $n = 2^7 \div 2^{11}$. Tai reiškia, kad tiek vaizdo V matematinių modelių seka (pav. 1), tiek orientuotas grafas-medis (pav. 2) yra baigtiniai. To priežastys akivaizdžios: iš vienos pusės, grafe-medyje, kaip taisyklė, sutinkamos viršūnės, neturinčios išeinančių iš jų lankų. Tai „galinės" grafo viršūnės $V_t(\tau)$, tenkinančios sąlygą: $\forall\lambda(\lambda \in \lambda_1\lambda_2 \cdots \mid \lambda_j \in 0\,1, 2, 3 \;\; j = 1, 2, \ldots \;\; |\lambda| \geqslant 1)$: viršūnių $V_{t+|\lambda|}(\tau\lambda)$ ir $V_t(\tau)$ skaitinės šviesos intensyvumo reikšmės sutampa (t-detalizacijos lygio modelyje V_t tokios „galinės" viršūnės atitinka toliau „nedalomus" blokelius, nusakomus adresu τ); iš kitos pusės, vaizdus apdorojančių, perduodančių ir priimančių įrenginių skiriamoji geba yra baigtinė.

Iš to kas buvo pasakyta išplaukia, jog vaizdą $V \in R$ bet kuriuo jo detalizacijos lygiu matematiškai galima aprašyti, nurodant visas adresų, vedančių į vaizdo V fragmentus (grafo-medžio viršūnes), kuriems priskirta viena ar kita skaitinė intensyvumo reikšmė, aibes. Reguliarios tokių adresų aibės nusakomos baigtiniais automatais (BA) arba reguliariosiomis išraiškomis [1]. BA teorijos panaudojimo, aprašant dvimačius realaus pasaulio vaizdus, idėja nėra nauja [2, 3, 4]. Šią idėją praplėtė ir pritaikė



Pav. 2.

vaizdų suspaudimui amerikiečiai K. Culik II ir J. Kari, ypatingą dėmesį skirdami „juodai baltiems" (siluetiniams) vaizdams [5, 6].

Žemiau pateikiama dvimačių vaizdų (su „pilka" šviesos intensyvumo skale) efektyvaus saugojimo (suspaudimo) metodika, pagrįsta BA teorijos bei vaizdo fragmentų tarpusavio panašumą didinančių procedūrų taikymu.

Pirmiausia įvesime dviejų vaizdų $V, W \in R$ to paties detalizacijos lygio s modelių panašumo (vidutinės kvadratinės paklaidos ε prasme) sampratą. Jeigu maksimali leistina vidutinės kvadratinės paklaidos (VKP) reikšmė lygi ε_0, tai modeliai (vaizdai) V_s ir W_s bus laikomi panašiais (žymėsime $V_s \sim W_s$) tada ir tik tada, kai

$$\varepsilon = \varepsilon\big(V_s, W_s\big) = \frac{1}{4^\varepsilon} \sum_{\omega:\ |\omega|=s} \big(W_s(\omega) - V_s(\omega)\big)^2 < \varepsilon_0. \tag{1}$$

Pastaba. Jeigu atliekama dviejų vaizdų, V_s ir W_t, priklausančių skirtingiems detalizacijos lygiams ($t > s$), panašumo analizė, tai (1) sąlygoje vietoj W_t imamas jo $(t - s)$ vienetais mažesnio detalizacijos lygio modelis (atitikmuo).

Vaizdų V_s ir W_s panašumo sąvoką galima ir tikslinga praplėsti tokiu būdu: prieš tikrinant (1) sąlygą, vaizdas V_s vienaip ar kitaip apdorojamas, t. y. jisai keičiamas vaizdu $\widetilde{V}_s$, gautu pritaikius kurią nors vieną (arba kelias) iš žemiau pateiktų transformacijų:

1) Vaizdas V_s dauginamas iš skaliaro $\mu > 0$, būtent: $\widetilde{V}_s = \mu V_s$, t.y. $\forall(|\omega| = s)$: $\widetilde{V}_s(\omega) = \mu V_s$; čia $\mu = (\sum_\omega V_s(\omega) \cdot W_s(\omega))/(\sum_\omega V_s^2(\omega))$; daugiklio μ įvedimas „priartina" vaizdus V_s ir W_s (VKP prasme);

2) Vaizdas V_s pasukamas $k \cdot 90°$ laipsnių kampu (pagal laikrodžio rodyklę; $k = 1, 2, 3$), t. y. palyginimui imamas ne V_s, o transformuotas vaizdas $\widetilde{V}_s$, kuriam teisinga: $\forall\omega(|\omega| = s)$: $\widetilde{V}_s(\omega) = V_s(\widetilde{\omega})$, kur $\widetilde{\omega} = \widetilde{\omega}_1\widetilde{\omega}_2 \cdots \widetilde{\omega}_s$, $\widetilde{\omega}_i = (\omega_i + k) \bmod 4$, $i = 1, 2, \ldots, s$;

3) Imamas vaizdas $\widetilde{V}_s$ gautas atlikus vaizdo (masyvo) V_s veidrodinį atspindį jo pagrindinės įstrižainės atžvilgiu bei posūkį $k \cdot 90°$ laipsnių kampu (prieš laikrodžio rodyklę; $k = 0, 1, 2, 3$). Šios transformacijos atlikimą garantuoja sąlygos išpildymas: $\forall\omega(|\omega| = s)$: $\widetilde{V}_s(\omega) = V_s(\widetilde{\omega})$, kur $\widetilde{\omega} = \widetilde{\omega}_1\widetilde{\omega}_2 \cdot \widetilde{\omega}_s$ ir $(\omega_i + \widetilde{\omega}_i) \bmod 4 = k$, $i = 1, 2, \ldots, s$.

Pastaba. 2 ir 3 punktuose nurodytoms vaizdo V_s transformacijoms tikslinga papildomai įvesti daugiklį μ (1 punktas);

4) Vaizdas V_s invertuojamas, t. y. $\forall\omega(|\omega| = s)$: $\widetilde{V}_s(\omega) = 256 - V_s(\omega)$. Gautam vaizdui papildomai galima taikyti bet kurią aukščiau paminėtą transformaciją (1, 2 arba 3 punktai).

Toliau pateikiama sudaryta BA konstravimo vaizdui $V \in R$ procedūra, minimizuojanti bendrą BA padėčių skaičių bei užtikrinanti efektyvaus vaizdo V aprašo (saugojimo, perdavimo prasme) gavimą.

Tarkime, jog duotas vaizdo V matematinis modelis V_n, $n = 2^7 \div 2^{11}$. Fiksuojama maksimali leistina VKP, naudojamos įvertinant atskirų vaizdą V sudarančių fragmentų panašumą, reikšmė ε_0. BA padėtį x atitinkantį vaizdą žymėsime $\nu(x)$.

1) $i = j = 0$.

2) Ivedame nulinę (pradinę) BA padėtį 0, ir $\nu(0) = V_n = [V_n(v)]$, $v = v_1 v_2 \cdots v_n$, $v_i = \{0, 1, 2, 3\}$; $i = 1, 2, \ldots, n$.

3) Tarkime, kad $\nu(t) = W_S = [W_S(\omega)] = [V_n(v')]$, kur $v' = v_1 v_2 \cdots v_{n-s}\omega$; čia $v_1 v_2 \cdots v_{n-s}$ yra fiksuotas $(n - s)$ ilgio adresas, o $\omega = \omega_1\omega_2 \cdots \omega_s$, $\omega_i \in \{0, 1, 2, 3\}$, $i = 1, 2, \ldots, s$; $0 < s < n$ t. y. vaizdas $\nu(i)$ prilygsta vienam iš 4^{n-s} galimų vaizdo V_n fragmentų $(2^s \times 2^s)$. Atskiru atveju, kai $s = n$, turime: $v_1 v_2 \cdots v_{n-s} = \varepsilon$, ir W_s sutampa su V_n (nulinė BA padėtis). BA padėtis i (vaizdas $[W_s(\omega)]$) apdorojama taip: kiekvienai k reikšmei iš aibės $\{0, 1, 2, 3\}$ tikrinama, ar yra bent viena jau įvesta BA padėtis q tokia, kuriai galiotų sąlyga: $\widetilde{W}_{s-1} \sim \nu(q)$, t. y. $\varepsilon(\nu(q), \widetilde{W}_{s-1}) \leqslant \varepsilon_0$; čia $W_{s-1} = [W_{s-1}(\omega')] = [W_s(k\omega')]$, $\omega' = \omega'_1\omega'_2 \cdots \omega'_{s-1}$, $\omega'_i \in \{0, 1, 2, 3\}$, $i = 1, 2, \ldots, s - 1$. Jei ši sąlyga išpildyta, tai fiksuojamas perėjimas iš padėties i į padėtį q, užrašant reikšmę k bei transformacijas, kurių pagalba buvo gautas vaizdas $\widetilde{W}_{s-1}$. Priešingu atveju, $j = j + 1$, $\nu(j) = W_{s-1}$ ir, įvedę naują BA padėtį j, reikšme k pažymime perėjimą iš i į j.

4) Jeigu $i = j$, t. y. visos BA padėtys (ir joms atitinkantys vaizdo V fragmentai) apdorotos, procedūrą baigiame. Priešingu atveju, $i = i + 1$, ir pereiname prie 3 punkto.

Pastaba. Praktiniu požiūriu netikslinga analizuoti vaizdo V fragmentus, kurių matmenys mažesni negu (8×8). Pastarųjų apdorojimui geriau taikyti kitas procedūras (BTC algoritmą, aproksimuojančias plokštumas ir panašiai [7, 8]).

Akivaizdu, kad sukonstruoto BA padėčių skaičius priklauso ne tiktai nuo pradinio vaizdo V detalizacijos lygio n, bet ir nuo pasirinktos VKP reikšmės ε_0.

Vaizdo V atkūrimui (tiksliau, jo įverčio gavimui) tuo pačiu detalizacijos lygiu galima panaudoti procedūrą, analogišką greitam dešifravimo algoritmui, aprašytam [5], įvertinus visų naujai įvestų vaizdo fragmentų panašumą didinančių transformacijų panaudojimą.

LITERATŪRA

[1] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley, 1979.

[2] J. Berstel and M. Morcrette, Compact representation of patterns by finite automata, *Proceedings Pixim'89*, Paris, 1989, 387–402.

[3] J. Berstel and A. A. Nait, Quadtrees generated by finite automata, *AFCET*, **61/62** (1989), 167–175.

[4] L. Staiger, Quadtrees and theHausdorf dimensions of pictures, *Workshop on Geometrical Problems of Image Processing*, Georgenthal, 1989, 173–178.

[5] K. Culik II and J. Kari, Image compression using weighted finite automata, *Computer and Graphics*, **17** (1993), 305–313.

[6] K. Culik II and J. Kari, Finite state methods for compression and manipulation of images, *Proceedings of Data Compression Conference 1995*, Snowbird, Utah, 1995, 142–151.

[7] P. Fränti, O. Nevalainen, and T. Kaukoranta, Compression of digital images by block truncation coding: A survey, *The Computer Journal*, **37** (4) (1994),

[8] J. Valantinas, *Dvimačių vaizdų blokinio segmentavimo ir kodavimo metodas*, Resp. konf. „Lietuvos mokslas ir pramonė" (sekcija „Informacinės technologijos") pranešimų medžiaga, Kaunas, 1997,

Finite automata based processing of two-dimensional images

J. Valantinas

Finite automata theory based methodics for the compression of two-dimensional gray-scale images is presented. Variable block-size segmentation procedure as well as techniques for the improvement of similarity of nonoverlapping parts of the image are applied. Compression effect is attained by producing an image modeling finite automaton with minimal number of states.