

Р. В. Маслій; А. Я. Кулик, д. т. н., проф.; С. С. Білошкурський

МЕТОД ВИЯВЛЕННЯ ОБЛИЧ НА КОЛЬОРОВОМУ ЗОБРАЖЕННІ

У статті запропоновано метод, що дозволяє виявити обличчя фронтального вигляду в діапазоні нахилів обличчя від центральної осі симетрії в межах від -56 до $+56$ градусів на кольоровому зображенні. Наведено результати дослідження ефективності виявлення обличчя на тестовому наборі зображень з бази Fddb.

Ключові слова: виявлення обличчя, Adaboost, кольорове зображення, модель кольору шкіри.

Вступ

Метод, що описаний в [1] дозволяє виявити обличчя фронтального вигляду в діапазоні нахилів обличчя від центральної осі симетрії в межах від -56 до $+56$ градусів на півтоновому зображенні. Перевагою методу в порівнянні з методом Віюлі – Джонса [2] є менша чутливість до змін освітлення, недоліком – більша обчислювальна складність. При обробці кольорових зображень зменшити обчислювальну складність при збереженні переваг методу можна шляхом використання кольору шкіри у якості ознаки для виявлення.

Сегментація зображення на пікселі «шкіри» та «не шкіри» потенційно дає можливість пришвидшити процес виявлення, так як дозволяє швидко відкинути області, в яких кількість пікселів «шкіри» менша порога, і подальше виявлення обличчя проводити в областях, в яких кількість пікселів «шкіри» більша порога.

У багатьох методах виявлення пікселів «шкіри» [3] дослідники намагаються дуже точно побудувати модель кольору шкіри для досягнення високого рівня вірогідності виявлення пікселів шкіри при низькому рівні хибного виявлення. Та при різних умовах освітлення значення пікселів можуть виходити за діапазони, передбачені моделлю, що призводить до зменшення точності виявлення пікселів шкіри і, в свою чергу, до зниження вірогідності виявлення обличчя та збільшення кількості хибних виявлень обличчя.

Для вирішення проблеми виходу значень пікселів за діапазони, передбачені моделлю кольору шкіри, при сегментації кольорових зображень доцільно використати модель кольору шкіри, яка замість визначення з високою точністю пікселів, що належать кольору шкіри, з високою точністю буде виключати пікселі, які не можуть належати кольору шкіри [4].

Мета досліджень – підвищення швидкості виявлення обличчя на кольорових зображеннях, які характеризуються широким діапазоном змін освітлення та орієнтацій обличчя, за умови зменшення кількості хибних виявлень та підвищення вірогідності їх виявлень.

Для досягнення поставленої мети необхідно розв'язати такі **задачі**:

1. Описати модель кольору шкіри для здійснення сегментації кольорового зображення на пікселі «шкіри» та «не шкіри».
2. Описати процес виявлення обличчя з допомогою робастного до змін освітлення boosting-методу [1].

Сегментація кольорового зображення

Нехай \overline{skin} та \overline{skin} – підпростори кольорового простору RGB, причому $RGB = \overline{skin} \cup \overline{skin}$ та $\overline{skin} \cap \overline{skin} = \emptyset$, тоді піксель у позиції (i,j) кольорового двовимірного зображення, визначений вектором $\langle R_{ij}, G_{ij}, B_{ij} \rangle$, який належить простору \overline{skin} є пікселем «шкіри», а піксель, який належить простору \overline{skin} є пікселем «не шкіри». Піксель $\langle R_{ij}, G_{ij}, B_{ij} \rangle$ належить до \overline{skin} , якщо він задовольняє одному чи кільком з таких правил:

$$(R_{ij}/G_{ij}) \notin [V_{min}^{RG}, V_{max}^{RG}], \quad (1)$$

$$(R_{ij}/B_{ij}) \notin [V_{min}^{RB}, V_{max}^{RB}], \quad (2)$$

$$(R_{ij} \in [0, V^R]) \cap (G_{ij} \in [0, V^G]) \cap (B_{ij} \in [0, V^B]), \quad (3)$$

де V_{min}^{RG} та V_{max}^{RG} – граничні значення відношення компоненти R до компоненти G , V_{min}^{RB} та V_{max}^{RB} – граничні значення відношення компоненти R до компоненти B , V^R , V^G та V^B – порогові значення для компонент R , G та B .

Описані у формулах (1 – 3) правила визначають модель кольору шкіри, після застосування якої до кольорового зображення, кожному пікселю $\langle R_{ij}, G_{ij}, B_{ij} \rangle \in \overline{skin}$ присвоюється значення одиниці, а кожному пікселю $\langle R_{ij}, G_{ij}, B_{ij} \rangle \in skin$ значення нуля, утворюючи таким чином бінарне зображення (бінарну маску).

Boosting-метод виявлення облич

Структурна схема роботи boosting-методу [1] при визначенні присутності в області півтонового зображення обличчя представлена на рис. 1.

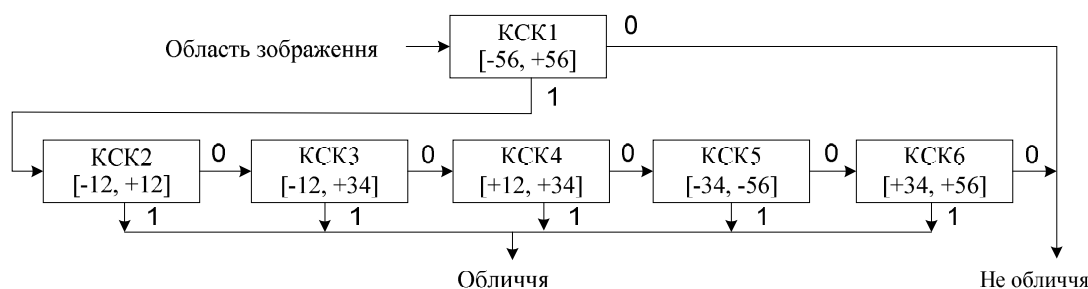


Рис. 1. Структурна схема роботи boosting-методу виявлення облич

Кожен з шести каскадів сильних класифікаторів ($KCK1$ – $KCK6$), зображених на рис. 1, тренується на окремих наборах навчальних зображень, які містять зображення обличчя у відповідному діапазоні нахилів, наприклад набір для $KCK2$ містить зображення обличчя фронтального вигляду з нахилами від центральної осі симетрії в діапазоні від -12° до $+12^\circ$. Кожний каскад сильних класифікаторів формується з використанням алгоритму навчання класифікації Adaboost. У якості ознак для простих класифікаторів у сильних класифікаторах використані модифіковані локальні бінарні шаблони ($RECT-LBP$). Використання цих ознак забезпечує високу робастність методу до змін умов освітлення у порівнянні з методами, в яких використовуються чутливі до змін освітлення хаароподібні ознаки.

Значення $RECT-LBP$ ознаки, після застосування $RECT-LBP$ до області зображення визначається таким чином:

$$RECT - LBP = \sum_{m=0}^7 s(k_m - k_c) 2^m, \quad (4)$$

де k_c – це сума значень інтенсивностей пікселів центрального прямокутника, k_m ($m=0, \dots, 7$) сума значень інтенсивностей пікселів сусідніх центральному прямокутників $s(x)$:

$$s(x) = \begin{cases} 1, & \text{якщо } x \geq 0 \\ 0, & \text{якщо } x < 0 \end{cases}.$$

Параметри dx та dy кожної $RECT-LBP$ ознаки визначають розміри центрального та сусідніх центральному прямокутників.

Під час формування кожного каскаду сильних класифікаторів з допомогою алгоритму Adaboost для кожної області зображення p усіх позитивних навчальних прикладів

обчислюється *RECT-LBP* ознака і формується масив $L_p^{\text{обличчя}}$, використовуючи набір негативних прикладів формується масив $L_p^{\text{не-обличчя}}$. Масиви $L_p^{\text{обличчя}}$ та $L_p^{\text{не-обличчя}}$ складаються з n_LBP елементів, причому кожний елемент відповідає певній *RECT-LBP* ознаці. В елементах масиву $L_p^{\text{обличчя}}$ додаються ваги позитивних прикладів, в елементах масиву $L_p^{\text{не-обличчя}}$ додаються ваги негативних прикладів. Вага прикладу додається до елемента відповідного масиву, якщо *RECT-LBP* ознака обчислена в області p відповідає *RECT-LBP* ознаці елемента. Помилка класифікації області p визначається за формулою:

$$\varepsilon_p = \sum_{j=1}^{n_LBP} \min(L_p^{\text{обличчя}}[j], L_p^{\text{не-обличчя}}[j]) \quad (5)$$

Формується простий класифікатор $h_t(x)$, використовуючи масиви $L_p^{\text{обличчя}}$ та $L_p^{\text{не-обличчя}}$ області p з мінімальною помилкою класифікації ε_t за формулою:

$$L_p[j] = \begin{cases} 1, & L_p^{\text{обличчя}}[j] > L_p^{\text{не-обличчя}}[j] \\ 0, & \text{інакше} \end{cases} \quad (6)$$

Формується сильний класифікатор за формулою:

$$h(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \psi \\ 0, & \text{інакше} \end{cases} \quad (7)$$

де α_t – коефіцієнт, значення якого визначається під час навчання, ψ – поріг, значення якого визначається при використанні верифікаційного набору прикладів.

За допомогою оптичного пристрою, наприклад відеокамери, зображення, на якому присутнє одне чи більше людських облич, вводять до комп'ютера. Це зображення перетворюють на інтегральне зображення. Мінімальне вікно пошуку, розмір якого збільшується на коефіцієнт Ms , при зміні масштабу, пересувається зображенням уздовж його рядків з кроком Kr пікселів, та виділяє область зображення. У випадку, якщо наповненість одиницями цієї області у бінарній масці не перевищує порога $NN\%$, область подається на вхід *KCK1*. В іншому випадку здійснюється перехід до обробки наступної області зображення. Якщо всі сильні класифікатори *KCK1* згідно формули (7) приймають значення $h(x) = 1$, то область зображення передається на вхід *KCK2*. Якщо ж хоча б один сильний класифікатор з *KCK1* прийме значення $h(x) = 0$, то детектор облич перейде до опрацювання наступної області вхідного зображення. Якщо хоча б один сильний класифікатор з *KCK2* – *KCK6* приймає значення $h(x) = 0$, то область зображення перейде на вхід наступного *KCK* або детектор перейде до опрацювання наступної області зображення, якщо поточним *KCK* був *KCK6*. Якщо ж всі сильні класифікатори одного з *KCK2* – *KCK6* прийняли значення $h(x) = 1$, то область зображення приймається за виявлену область-кандидат в діапазоні кутів нахилу поточного *KCK*.

Після проходження вікна пошуку по всьому зображенню у всіх масштабах детектор обличчя для кожного діапазону кутів нахилу виділяє на зображенні декілька областей-кандидатів. Пропонується застосувати наступний етап верифікації областей-кандидатів. У межах кожного діапазону кутів нахилу застосовується правило кластеризації, в якому уся множина областей-кандидатів розбивається на множини, що не перетинаються. При кластеризації дві області-кандидати відносяться до одного кластера, якщо вони перетинаються більше ніж на відсоток φ . Кластер є кандидатом для створення об'єднаної області, якщо множина областей-кандидатів у ньому більша порога η . Об'єднана область формується шляхом знаходження середнього арифметичного координат областей-

кандидатів, що входять до кластера. Для формування областей облич між усіма діапазонами кутів нахилу використовується показник якості w , який визначається кількістю областей-кандидатів, які увійшли до об'єднаної області. Якщо об'єднані області різних діапазонів кутів нахилу перетинаються більш ніж на μ відсотків областю обличчя визнається об'єднана область з більшим показником якості. Якщо об'єднані області різних діапазонів кутів нахилу перетинаються менше ніж на μ відсотків, то кожна з об'єднаних областей визнається як область обличчя.

Експериментальні дослідження методу

Метод, що пропонується, складається з двох етапів. На першому здійснюється сегментація кольорового зображення на пікселі «шкіри» та «не шкіри» при використанні моделі кольору шкіри, визначеної правилами (1 – 3), створюється бінарна маска, кольорове зображення перетворюється на півтонове. На другому етапі області півтонового зображення обробляються за схемою, представленою на рис. 1, область буде подаватися на вхід $KCKI$ у випадку, якщо наповненість одиницями цієї області у бінарній масці не перевищує $NN\%$.

Для дослідження методу були використані стандарт та база зображень облич FDDDB [4]. Зображення у цій базі характеризуються широким діапазоном змін: у орієнтації облич, фоні, освітленні та зовнішньому вигляді. Стандарт містить анотації кожного обличчя у вигляді еліптичних областей, а також набір програм для оцінки ефективності виявлення.

Під час дослідження сегментації за кольором шкіри на тестовому наборі з 200 кольорових зображень були визначені такі параметри моделі кольору шкіри: $V_{\min}^{RG} = 1,1$, $V_{\max}^{RG} = 3,1$, $V_{\min}^{RB} = 1,04$, $V_{\max}^{RB} = 5$, $V^R = 55$, $V^G = 35$ та $V^B = 25$. Приклад сегментації кольорових зображень з бази FDDDB на пікселі «шкіри» та «не шкіри» представлений на рис. 2.



Рис. 2. Приклад сегментації кольорових зображень з бази FDDDB на пікселі «шкіри» та «не шкіри»: а – в) оригінальні кольорові зображення; г – е) бінарні маски, отримані при застосуванні моделі кольору шкіри

Дослідження методу здійснювалося при таких параметрах: кількість вікон $\eta = 2$, коефіцієнт масштабу $Ms = 1.2$, розмір мінімального вікна пошуку (48×48 пікселів), розмір ядра $RECT-LBP - 3 \times 3$ пікселів, $n_LBP = 256$, зсув вікна пошуку $Kr = 1$ піксель, відсоток перетину вікон $\varphi = 70\%$, поріг $\eta = 2$, показник якості $w = 2$, відсоток перетину вікон $\mu = 60\%$, наповненість одиницями вікна пошуку у бінарній масці $NN = 60\%$.

Для оцінки ефективності виявлення був використаний тестовий набір зі 120 кольорових зображень при використанні комп'ютера AMD 3000+ з частотою процесора 1,81 ГГц. Результати дослідження розробленого методу і boosting-методу [1] представлені в таблиці 1.

Таблиця 1

Результати дослідження методів виявлення облич

Критерії оцінки ефективності виявлення	Boosting-метод без сегментації за кольором шкіри [1]	Розроблений метод
Вірогідність виявлення, %	76,4	75,9
Кількість хибних виявлень, шт.	38	31
Швидкість виявлення, кадрів/сек.	2 – 4	3 – 6

Висновки

Розроблений метод дозволяє підвищити швидкість обробки зображень у порівнянні з boosting-методом [1]. Досягаючи при цьому зменшення на 18% кількості хибних виявлень при незначному зменшенні вірогідності виявлення на тестовому наборі з бази FDDB.

СПИСОК ЛІТЕРАТУРИ

1. Патент № 53412 Україна, МПК (2006) G06K 9/62. Спосіб знаходження людських облич на зображенні / Маслій Р. В., Кулик А. Я.; завник та патентовласник Вінницький національний технічний університет. – № u201002853; заявл. 15.03.2010; опубл. 11.10.2010, Бюл. № 19.
2. Viola P. Robust Real-Time Face Detection / P. Viola, M. Jones // International Journal of Computer Vision. – 2004. – Vol. 57, № 2. – P. 137 – 154.
3. Vezhnevets V. A Survey on Pixel-Based Skin Color Detection Techniques / V. Vezhnevets, V. Sazonov, A. Andreeva // Proc. Graphicon-2003, Moscow, Russia, September 2003, P. 85 – 92.
4. Маслій Р. В. Метод виявлення облич на кольоровому зображенні / Р. В. Маслій, А. Я. Кулик, С. С. Білошкурський // Контроль і управління в складних системах (КУСС-2010). Тези доповідей десятої міжнародної науково-технічної конференції. – 2010. – С. 144.
5. Vidity Jain and Erik Learned-Miller. FDDB: A Benchmark for Face Detection in Unconstrained Setting. Technical Report UM-CS-2010-009, Dept. of Computer Science, University of Massachusetts, Amherst. 2010 [Електронний ресурс] / Режим доступу: <http://vis-www.cs.umass.edu/fddb/fddb.pdf>.

Маслій Роман Васильович – асистент кафедри автоматики та інформаційно-вимірювальної техніки

Кулик Анатолій Ярославович – д. т. н., професор кафедри автоматики та інформаційно-вимірювальної техніки.

Білошкурський Сергій Степанович – студент кафедри автоматики та інформаційно-вимірювальної техніки.

Вінницький національний технічний університет.