



УКРАЇНА

(19) **UA** (11) **54760** (13) **U**
(51) МПК (2009)
G06K 9/62

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ

ОПИС ДО ПАТЕНТУ НА КОРИСНУ МОДЕЛЬ

видається під
відповідальність
власника
патенту

(54) СПОСІБ ВИЯВЛЕННЯ ОБЛИЧ ЛЮДЕЙ НА ЗОБРАЖЕННІ

1

2

(21) u201005393

(22) 05.05.2010

(24) 25.11.2010

(46) 25.11.2010, Бюл.№ 22, 2010 р.

(72) МАСЛІЙ РОМАН ВАСИЛЬОВИЧ, КУЛИК АНАТОЛІЙ ЯРОСЛАВОВИЧ, БІЛОШКУРСЬКИЙ СЕРГІЙ СТЕПАНОВИЧ

(73) ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

(57) Спосіб виявлення облич людей на зображенні, в якому формують базу зображень облич фронтального вигляду та базу зображень, на яких відсутні обличчя, з яких формують навчальні та верифікаційні набори позитивних та негативних прикладів, тренують каскад сильних класифікаторів, додають до каскаду сильні класифікатори до тих пір, поки не буде досягнуто заданих значень помилки першого роду та рівня детекції для каскаду; тренують сильний класифікатор, додають слабкі класифікатори у сильний класифікатор до тих пір, поки не буде досягнуто заданих значень помилки першого роду та рівня детекції для сильного класифікатора, тренують слабкі класифікатори, додають у сильний класифікатор слабкий класифікатор, який на зваженому наборі навчальних прикладів показав найменшу помилку, формують сильний класифікатор, на верифікаційних наборах перевіряють досягнення помилки першого роду та рівня детекції для сильного класифікатора, негативний навчальний набір наступного класифікатора формують з прикладів, отриманих з бази зображень, на яких відсутні обличчя, які попередні

сильні класифікатори каскаду прийняли за обличчя, вводять у комп'ютер зображення, за допомогою вікна пошуку у різних масштабах і положеннях виділяють області зображення, подають їх на вхід каскаду сильних класифікаторів, область зображення признають за детектовану область, якщо всі сильні класифікатори каскаду прийняли область зображення за обличчя, кластеризують детектовані області в об'єднані області, кожна з яких признається за область обличчя, який **відрізняється** тим, що значення порога та парності ознак у слабких класифікаторах обчислюється до навчання сильного класифікатора, для формування слабких класифікаторів використовують модифіковані Хаар-подібні ознаки, кожна з яких характеризується значенням локального бінарного шаблону, обчислюють значення модифікованої Хаар-подібної ознаки шляхом підрахунку кількості пікселів у темних та світлих прямокутниках ознаки, в яких значення локального бінарного шаблону, застосованого до околу пікселя, відповідають значенню локального бінарного шаблону, яким характеризується ознака, для обчислення значень модифікованої Хаар-подібної ознаки використовують масив значень локальних бінарних шаблонів, який утворений шляхом застосування локального бінарного шаблону до пікселів півтонового зображення, а також масив інтегральних значень локальних бінарних шаблонів, негативні та позитивні навчальні набори являють собою масиви значень локальних бінарних шаблонів.

Корисна модель відноситься до автоматичної та обчислювальної техніки і може бути використана для розроблення комп'ютерних систем ідентифікації, систем інтелектуальної взаємодії "людина-комп'ютер", систем теле- та відеоконференцій, автоматичних систем слідування і безпеки.

Спосіб комп'ютерного розпізнавання людини за зображенням її обличчя (Патент UA №87181 опубл. в 2009р., М.кл. G06K9/62), що полягає у введенні зображення в комп'ютер, масштабуванні зображення, обробці його з виявленням фрагментів й опорних точок зображення, повороті зображення в необхідне положення, почерговому на-

(19) **UA** (11) **54760** (13) **U**

кладенні на оброблене зображення шаблонів, на введеному в комп'ютер зображенні спочатку знаходять характерні точки обличчя людини, за взаємним розташуванням яких визначають кути нахилу й повороту голови на зображенні, формують об'ємну модель середнього обличчя, масштабують її, накладають модель, повернену відповідно до певних кутів нахилу й повороту голови, на зображення, сполучають всі точки моделі з пікселями вихідного зображення, будують фронтальний вид введеного зображення обличчя й розпізнають людину.

Недоліком є те, що у відомому способі комп'ютерного розпізнавання людини за зображенням її обличчя у випадку змін освітлення чи використання зображень низької якості характерні точки обличчя при кореляційному порівнянні шаблонів цих точок можуть бути не знайдені, що може призвести до збільшення помилок другого роду.

Найбільш близьким до способу, що заявляється, є спосіб знаходження об'єктів на зображенні (Патент US 7099510, М.кл. G06K9/62, опубл. в 2006р.), в якому спочатку формують базу зображень обличчя фронтального вигляду V_{face} , та базу зображень, на яких відсутні обличчя V_{nonface} . Зображення з V_{face} нормалізують таким чином, щоб очі на зображеннях знаходилися в одних і тих же просторових координатах та формують з них навчальний P та верифікаційний V_p набори позитивних прикладів.

Виділяють з V_{nonface} області зображень та формують з них навчальний N та верифікаційний V_n набори негативних прикладів.

Зображення у наборах P, V_p, N, V_n приводять до розміру 24×24 пікселі. Визначають набір прямокутних Хаар-подібних ознак, кожна з яких визначається типом, розмірами та місцезнаходженням на зображенні розміром 24 на 24 пікселі.

Потім тренують каскад сильних класифікаторів (КСК), для чого встановлюють значення помилки першого роду FP_C , рівень детекції TP_C яких при тренуванні має досягнути КСК. Додають до КСК сильні класифікатори (СК) до тих пір, поки не будуть досягнуті встановлені FP_C та TP_C . Встановлюють значення помилки першого роду FP , та рівня детекції TP для кожного СК. Кожен СК тренують до досягнення встановлених FP та TP наступним чином. Формують навчальну вибірку з набору P позитивних прикладів та набору N негативних прикладів вигляду $(x_1, y_1), \dots, (x_n, y_n)$, де x - зображення, y - цільове значення, яке рівне 0 - для негативних прикладів, 1 - для позитивних прикладів. Ініціалізують ваги $w_{1,i}=1/2m$, $1/2l$, для $y_i=0$, 1 відповідно, де m - кількість позитивних прикладів, l - кількість негативних прикладів, а поточний рівень помилки першого роду $fpRate=1$. Запускають цикл з умовою поки $fpRate > FP$, де t - це поточна ітерація, i - індекс прикладу, j - індекс Хаар-подібної ознаки. Нормалізують ваги прикладів навчальної вибірки: $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{k=1}^n w_{t,k}}$. Для кожної Хаар-подібної

ознаки f_j обчислюють її значення для всіх прикладів навчальної вибірки, визначають оптимальний поріг θ_j , та напрямок знаку p_j . Формують слабкий класифікатор (СлК) для кожної Хаар-подібної ознаки за формулою:

$$h_j(x) = \begin{cases} 1, & \text{якщо } p_j f_j(x) < p_j \theta_j \\ 0, & \text{інакше} \end{cases}$$

Обчислюють помилку усіх СлК на навчальній вибірці за формулою: $\varepsilon_j = \sum_i w_i |h_i(x_i) - y_i|$. У СК додають СлК h_t з найменшою помилкою ε_t . Оно-влюють ваги прикладів за формулою: $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$, де $e_i=0$ якщо приклад x_i розпізнаний правильно, $e_i=1$ в інакшому випадку,

$$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}.$$

Формують СК:

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

$$\text{Де } \alpha_t = \log \frac{1}{\beta_t}, \psi = \frac{1}{2} \sum_{t=1}^T \alpha_t.$$

Використовуючи верифікаційний набір позитивних прикладів V_p підлаштовують поріг ψ СК, для досягнення заданого рівня детекції TP . Обчислюють помилку першого роду $fpRate$ СК на верифікаційному наборі негативних прикладів V_n . Переходять на наступну ітерацію циклу у випадку виконання умови $fpRate > FP$, в іншому випадку СК додають у КСК.

Переходять до тренування наступного СК. Для формування набору негативних прикладів TU наступного СК виділяють з V_{nonface} області зображень, приводять до розміру 24 на 24 пікселя формують з них негативні приклади і подають на вхід КСК, приклади, які КСК прийняв за позитивні додають у набір N .

За допомогою оптичного пристрою, наприклад, відеокамери, зображення, на якому присутнє одне чи більше людських обличчя вводять у комп'ютер. Це зображення перетворюють у інтегральне зображення. За допомогою віка пошуку, розмір якого спочатку дорівнює 24×24 пікселі, а потім збільшують на коефіцієнт M_s , при зміні масштабу, пересувають по зображенню уздовж його рядків з кроком K_g пікселів, виділяють область зображення, яку подають на вхід КСК. У СК з допомогою інтегрального зображення нормалізують виділені області та обчислюють значення прямокутних Хаар-подібних ознак у СлК. Згідно формули (1) приймають рішення про перехід до наступного СК у каскаді у випадку $h(x)=1$, чи до відхилення поточної області зображення у випадку $h(x)=0$ і переходу до обробки наступної області зображення виділеної вікном пошуку. У випадку якщо всі СК каскаду приймуть рішення $h(x)=1$ область зображення признається за таку, на якій присутнє обличчя.

Після проходження вікна пошуку по всьому зображенню у всіх масштабах детектор обличчя виділяє на зображенні декілька детектованих областей. Застосовують правило кластеризації, в якому всю множину детектованих областей розбивають на множини, що не перетинаються. При кластеризації дві детектовані області відносять до одного кластеру, якщо вони перетинаються більше ніж на відсоток ϕ . Кластер є кандидатом для створення об'єднаної області, якщо множина детектованих областей у ньому більша порогу η . Об'єднану область формують шляхом знаходження середнього арифметичного координат детектованих областей, що входять у кластер. Кожну з об'єднаних областей признають за область обличчя.

Недоліком є те, що у відомому способі знаходження об'єктів на зображенні застосовані у якості СЛК Хаар-подібні ознаки чутливі до змін освітлення, тому перед знаходженням обличчя потрібно виконувати спеціальну процедуру корекції освітлення зображення. Крім того етап навчання детектора є досить тривалим.

В основу корисної моделі поставлена задача створення такого способу виявлення обличчя людей на зображенні, який забезпечить надійне та швидке (режим реального часу) виявлення обличчя, чи декількох обличчя у фронтальному ракурсі, в умовах змін освітлення та масштабу без використання дорогого спеціалізованого обладнання для визначення просторових координат обличчя.

Поставлена задача досягається тим, що в способі виявлення обличчя людей на зображенні, в якому формують базу зображень обличчя фронтального вигляду та базу зображень, на яких відсутні обличчя, з яких формують навчальні та верифікаційні набори позитивних та негативних прикладів, тренують каскад сильних класифікаторів, додають до каскаду сильні класифікатори до тих пір, поки не буде досягнуто заданих значень помилки першого роду та рівня детекції для каскаду; тренують сильний класифікатор, додають слабкі класифікатори у сильний класифікатор до тих пір, поки не буде досягнуто заданих значень помилки першого роду та рівня детекції для сильного класифікатора, тренують слабкі класифікатори, додають у сильний класифікатор слабкий класифікатор, який на зваженому наборі навчальних прикладів показав найменшу помилку, формують сильний класифікатор, на верифікаційних наборах перевіряють досягнення помилки першого роду та рівня детекції для сильного класифікатора, негативний навчальний набір наступного класифікатора формують з прикладів, отриманих з бази зображень, на яких відсутні обличчя, які попередні сильні класифікатори каскаду прийняли за обличчя, вводять у комп'ютер зображення, за допомогою вікна пошуку у різних масштабах і положеннях виділяють області зображення, подають їх на вхід каскаду сильних класифікаторів, область зображення признають за детектовану область, якщо всі сильні класифікатори каскаду прийняли область зображення за обличчя, кластеризують детектовані області в об'єднані області, кожна з яких признається за область обличчя, причому значення порогу та парності ознак у слабких класифікаторах обчислю-

ють до навчання сильного класифікатора, для формування слабких класифікаторів використовують модифіковані Хаар-подібні ознаки, кожна з яких характеризується значенням локального бінарного шаблону, обчислюють значення модифікованої Хаар-подібної ознаки шляхом підрахунку кількості пікселів, в яких значення локального бінарного шаблону застосованого до околу пікселя відповідають значенню локального бінарного шаблону, яким характеризується ознака, для обчислення значень модифікованої Хаар-подібної ознаки використовують масив значень локальних бінарних шаблонів, який утворений шляхом застосування локального бінарного шаблону до пікселів півтонового зображення, а також масив інтегральних значень локальних бінарних шаблонів, негативні та позитивні навчальні набори представляють собою масиви значень локальних бінарних шаблонів.

На Фіг.1 представлені типи модифікованих Хаар-подібних ознак (МХО), на Фіг.2 показаний приклад обчислення значення локального бінарного шаблону CS-LBP з параметрами: кількість точок околу $N=8$, радіус околу $R=1$, поріг $\lambda=0$ після згортки з областю півтонового зображення. На Фіг.3 показаний приклад отримання масиву значень локальних бінарних шаблонів (масив ЛБШ) (розміром 3×3) з півтонового зображення (розміром 5×5 пікселів) використовуючи CS-LBP, на Фіг.4. показаний приклад обчислення масиву інтегральних значень локальних бінарних шаблонів (масив ІЛБШ) з масиву ЛБШ (розміром 3×3), при різних параметрах d , на Фіг.5 наведений алгоритм прийняття рішення КСК щодо присутності на області зображення обличчя.

Спосіб здійснюється наступним чином. Спочатку формують базу зображень обличчя фронтального вигляду V_{face} , та базу зображень, на яких відсутні обличчя $V_{nonface}$. Зображення з V_{face} нормалізують таким чином, щоб очі на зображеннях знаходилися в одних і тих же просторових координатах та формують з них навчальний P та верифікаційний V_p набори позитивних прикладів.

Виділяють з $V_{nonface}$ області зображень та формують з них навчальний N та верифікаційний V_n набори негативних прикладів. Зображення у наборах P, V_p, N, V_n приводять до розміру 24×24 пікселі.

Визначають набір МХО, з такими параметрами: тип, значення CS-LBP, розміри та місцезнаходження на зображенні розміром 24 на 24 пікселі. Значення CS-LBP ознаки після застосування CS-LBP до околу пікселя зображення з координатами (x,y) визначається наступним чином:

$$CS-LBP(x,y) = \sum_{m=0}^{(N/2)-1} s \left(k_m - k_{m+N/2} \right) z^m, \text{ де } N -$$

кількість пікселів околу, k - значення інтенсивностей пікселів околу,

$s_{k_m - k_{m+N/2}} = \begin{cases} 1, \text{ якщо } k_m - k_{m+N/2} > \lambda, \\ 0, \text{ інакше} \end{cases}$, де λ - поріг CS-LBP ознаки.

Потім тренують каскад сильних класифікаторів (КСК), для чого встановлюють значення помилки першого роду FP_C, рівень детекції TP_C яких при тренуванні має досягнути КСК. Додають до КСК СК до тих пір, поки не будуть досягнуті встановлені FP_C та TP_C.

З набору P отримують набір масивів ЛБШ позитивних прикладів I_{LBP}^P шляхом обчислення у пікселях зображень набору P CS-LBP ознак. Таким чином масиви ЛБШ з набору I_{LBP}^P будуть складатися зі значень CS-LBP ознак. З набору I_{LBP}^P отримують набір масивів ЛБШ негативних прикладів I_{LBP}^N використовуючи формули:

$$i_d(x, y) = i_d(x, y - 1) + \delta_d(x, y), \quad (2)$$

$$I_d(x, y) = I_d(x, y - 1, y) + i_d(x, y), \quad (3)$$

де I - масив ЛБШ, i - допоміжний масив, $d = 1, \dots, 16$; $\delta_d(x, y) = 1$, якщо комірка масиву з набору I_{LBP}^P з координатами x, y дорівнює d , і $\delta_d(x, y) = 0$ в інакшому випадку; $x, y = 1, \dots, 24$; нульові елементи масивів i та I (в яких $x=0$ або $y=0$) заповнюють нулями.

З набору N формують набір масивів ЛБШ негативних прикладів I_{LBP}^N . З набору I_{LBP}^N отримують набір масивів ЛБШ негативних прикладів I_{LBP}^N .

Використовуючи набори I_{LBP}^P та I_{LBP}^N отримують значення кожної МХО f_i та визначають її оптимальний поріг θ_j , та напрямок знаку r_j .

Значення МХО обчислюють наступним чином. Відповідно до параметрів МХО використовуючи відповідний масив ЛБШ I обчислюють кількість CS-LBP ознак у світлих та темних прямокутниках МХО. Значення МХО обчислюють за формулою: $f(I) = S_w - S_b$, де S_w - кількість CS-LBP ознак у світлих прямокутниках МХО, S_b - кількість CS-LBP ознак у темних прямокутниках МХО. Значення S_w та S_b обчислюють за формулою:

$$S = I(a_2, b_2) - I(a_3, b_3) - I(a_1, b_1) + I(a_4, b_4), \quad (4)$$

де (a_1, b_1) , (a_2, b_2) , (a_3, b_3) , (a_4, b_4) - це координати чотирьох пікселів пов'язаних з відповідним прямокутником у МХО (сусідній зверху з правим верхнім пікселем прямокутника, правий нижній піксель прямокутника, сусідній зліва з лівим нижнім пікселем прямокутника, сусідній по діагоналі зверху зліва з лівим верхнім пікселем прямокутника).

Встановлюють значення помилки першого роду FP, та рівня TP для кожного СК. Кожен СК тренують до досягнення встановлених FP та TP наступним чином.

Формують навчальну вибірку з набору I_{LBP}^P позитивних прикладів та набору I_{LBP}^N негативних прикладів вигляду $(x_1, y_1), \dots, (x_n, y_n)$, де x - масив ЛБШ, y - цільове значення, яке рівне 0 - для негативних прикладів, 1 - для позитивних прикладів. Ініціалізують ваги $w_{1,i} = 1/2m$, $1/2l$, для $y_i = 0, 1$ відповідно, де m - кількість позитивних прикладів, l - кількість негативних прикладів, а поточний рівень помилки першого роду fpRate=1. Запускають цикл з умовою поки fpRate > FP, де t - це поточна ітерація, i - індекс прикладу, j - індекс МХО.

Нормалізують ваги прикладів навчальної вибі-

$$\text{рки: } w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{k=1}^n w_{t,k}}.$$

Формують СлК h_j для кожної МХО використовуючи I_{LBP}^P та I_{LBP}^N за формулою:

$$h_j(x) = \begin{cases} 1, \text{ якщо } r_j f_j < p_j \theta_j \\ 0, \text{ інакше} \end{cases}$$

Де f_j - МХО, x - масив ЛБШ, θ_j - поріг, r_j - напрямок знаку МХО.

Обчислюють помилку усіх СлК на навчальній вибірці за формулою: $\varepsilon_j = \sum_i w_i |h_j(x_i) - y_i|$. У СК додають СлК h_t з найменшою помилкою ε_t . Оно-влюють ваги прикладів за формулою: $w_{t+1,i} = w_{t,i} \beta_t^{1-\varepsilon_i}$, де $\varepsilon_i = 0$ якщо приклад x_i розпізнаний правильно, $\varepsilon_i = 1$ в інакшому випадку,

$$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}. \text{ Формують СК:}$$

$$h(x) = \begin{cases} 1, \sum_{t=1}^T \alpha_t h_t(x) \geq \psi \\ 0, \text{ інакше} \end{cases} \quad (5)$$

$$\text{де } \alpha_t = \log \frac{1}{\beta_t}, \psi = \frac{1}{2} \sum_{t=1}^T \alpha_t.$$

Використовуючи верифікаційний набір позитивних прикладів V_p підлаштовують поріг ψ СК, для досягнення заданого рівня детекції TP. Обчислюють помилку першого роду fpRate СК на верифікаційному наборі негативних прикладів V_n . Переходять на наступну ітерацію циклу у випадку виконання умови fpRate > FP, в іншому випадку СК додають у КСК.

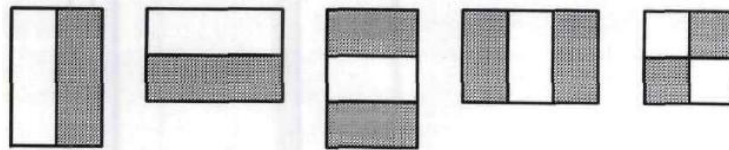
Переходять до тренування наступного СК. Для формування набору негативних прикладів N наступного СК виділяють з V_{nonface} області зображень, приводять до розміру 24 на 24 пікселя формують з них негативні приклади, перетворюють їх у масиви ЛБШ і подають на вхід КСК, приклади, які КСК прийняв за позитивні додають у набір N.

За допомогою оптичного пристрою, наприклад відеокамери, зображення на якому присутнє одне чи більше людських обличчя вводять у комп'ютер. Це зображення перетворюють у масив ЛБШ та з допомогою формул (2) та (3) створюють 16 масивів ЛБШ. За допомогою вікна пошуку, розмір якого спочатку дорівнює 24*24 пікселі, а потім збільшують на коефіцієнт Ms, при зміні масштабу, пересувають по зображенню уздовж його рядків з кроком Kp пікселів, виділяють область зображення, відповідні її частини масивів ЛБШ подають на вхід КСК. У СК використовуючи отримані частини масивів ЛБШ за формулою (4) обчислюють значення МХО у СлК. Згідно формули (5) приймають рішення про перехід до наступного СК у каскаді у випадку $h(x)=1$, чи до відхилення поточної області зображення у випадку $h(x)=0$ і переходу до обробки наступної області зображення виділеної вікном

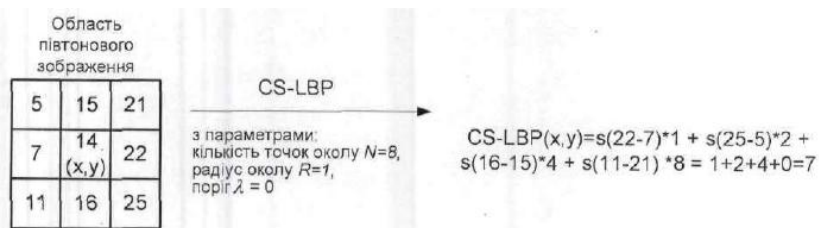
пошуку. У випадку якщо всі СК каскаду приймуть рішення $h(x)=1$ область зображення признається за таку, на якій присутнє обличчя.

Після проходження вікна пошуку по всьому зображенню у всіх масштабах детектор обличчя виділяє на зображенні декілька детектованих областей. Застосовують правило кластеризації, в якому всю множину детектованих областей розбивають на множини, що не перетинаються. При кластеризації дві детектовані області відносять до одного

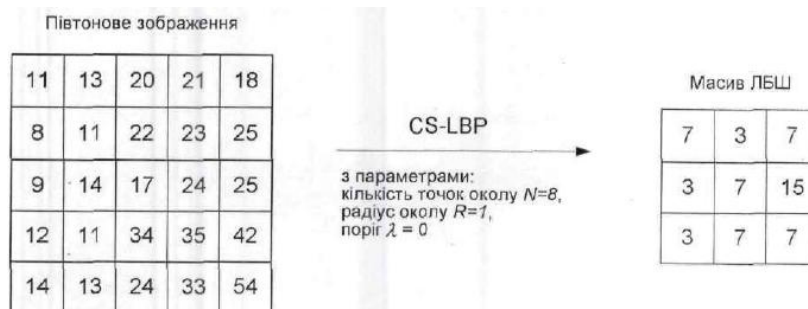
кластеру, якщо вони перетинаються більше ніж на відсоток Φ . Кластер є кандидатом для створення об'єднаної області, якщо множина детектованих областей у ньому більша порогу η . Об'єднану область формують шляхом знаходження середнього арифметичного координат детектованих областей, що входять у кластер. Кожну з об'єднаних областей признають за область обличчя.



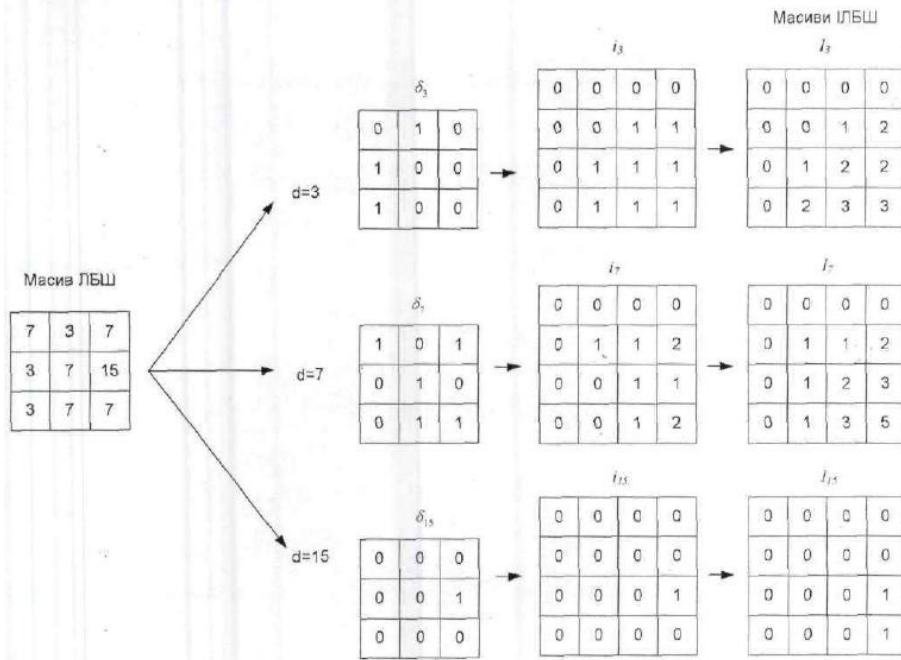
Фіг. 1



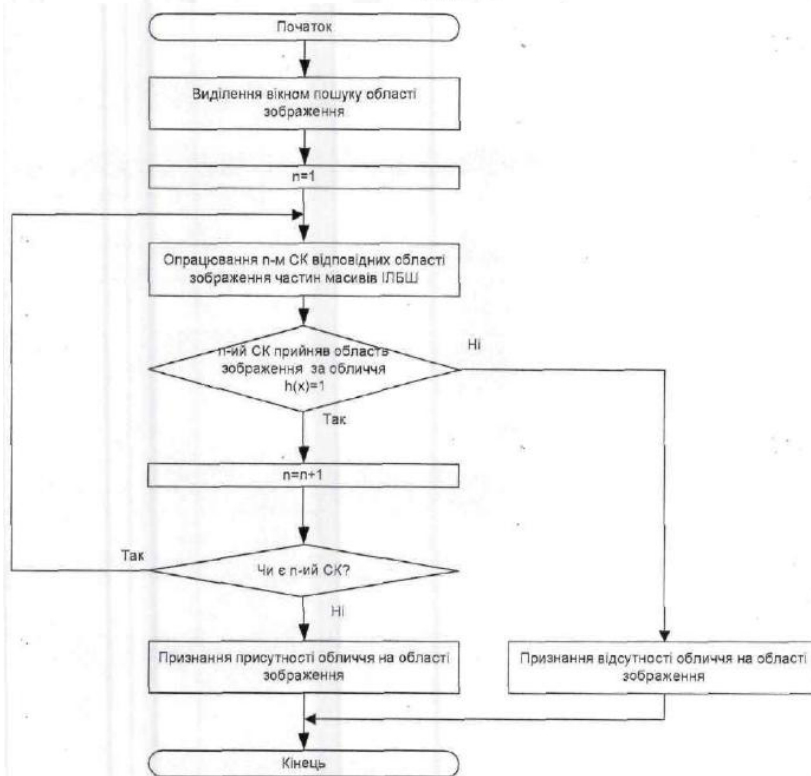
Фіг. 2



Фіг. 3



Фиг. 4



Фиг. 5