

Використання елементів довжини у структурному аналізі контурів об'єктів бінарних зображень.

Володимир Калмиков, Віталій Вишневський, Дмитро Калмиков

Інститут проблем математичних машин і систем НАН України,

252187 Київ, проспект академіка Глушкова 42

Тел. (044) 266 - 4147 E-mail kvg@immsp.kiev.ua

Abstract

The new method of the structural analysis of the binary image contours is proposed. The special structural elements is used, permitting to reduce the two-dimensional task of the contour analysis to a number of the one-dimensional tasks.

ВСТУП

Виконання цієї роботи обумовлене необхідністю обробки, зберігання та передачі швидко зростаючих об'ємів візуальної інформації, зокрема такої, як контури об'єктів на зображеннях медичних препаратів, інше. Цій проблемі останнім часом присвячено багато робіт, зокрема, структурному аналізу контурів об'єктів на зображеннях, наприклад [4,5]. На наш погляд прогресивними вбачаються методи структурного аналізу контурів [1,2,3,4], як такі, що використовують поняття характерних точок контура. Під характерними розуміються точки сполучення ліній, що утворюють контур об'єкту.

ЕЛЕМЕНТИ ДОВЖИНИ

як структурні елементи контуру.

Перевагами таких методів аналізу контурів вважають можливість вирішення задач обробки та розпізнавання під час одного проходу, суттєва економія пам'яті, що потрібна для зберігання зображення, і, як наслідок, зниження часу, потрібного для вирішення задач обробки зображення. У той же час лінії що розташовані під кутом близько 45° до ліній ґратки потребують занадто багато характерних точок, так багато, що іноді навіть зводить нанівець отриману економію пам'яті. Велика кількість характерних точок, у такому випадку, утруднює обробку. Взагалі природно було б у якості структурних елементів контуру визначити відрізки прямих та дуги довільних кривих ліній. Створення методу, який визначає довільний контур, як послідовність відрізків прямих та дуг довільних кривих ліній можна вважати б дещо віддаленою метою, у той час як першим кроком доцільно визначити більш економічний засіб відображення прямих ліній довільного напрямку на прямокутній ґратці. Для цього необхідно визначити додатково проміжні напрями до тих чотирьох, що на

цей час традиційно використовуються для опису ліній контуру: "уверх", "праворуч", "униз", "ліворуч". Всякий новий напрямок на ґратці може бути відображений тільки парами граней. Досить очевидно впливає з визначення алгоритму дискретизації, що при відображенні ліній, що описуються безперервними функціями, на ґратці довжина граней по меншій мірі одного з напрямів, що чергуються - вертикального чи горизонтального - дорівнює одиниці. Дійсно, зважаючи на те, що усяка функція відбивається при дискретизації [6] з використанням найменших можливих прирощень, мінімально можливі прирощення аргументу x та будь якої безперервної функції $y = f(x)$ при відображенні на ґратці $\Delta x = \Delta y = 1$. У тому випадку, якщо перша похідна функції $y = f(x)$ - $|\partial y / \partial x| \geq 1$, вертикальна грань $l_y = |(\partial y / \partial x) \cdot \Delta x| \geq 1$, у той час як $l_x = |\Delta x| = 1$. Якщо перша похідна функції $y = f(x)$ перебуває у межах $0 < |\partial y / \partial x| < 1$, горизонтальна грань $l_x = |(\partial y / \partial x) \cdot \Delta x| \geq 1$, у той час як $l_y = |\Delta y| = 1$

Сполучені пари взаємно перпендикулярних відрізків - *L-елементи* відповідають проміжним відносно тих, що були перераховані, напрямкам і здатні взяти на себе функцію елементів довжини - складових контуру. Вони утворюються парами граней, що розташовані одна за одною - вертикальною та горизонтальною, чи навпаки. Перша грань може мати довільну довжину: $l = 1, 2, \dots, L$, де L - гранично можлива на даній ґратці довжина, у той час як довжина другого сполученого відрізка дорівнює 1 (у деяких випадках дорівнює 0). Отже параметрами *L-елементу* є: l - довжина першої грані, t - напрям - один з основних, перелікованих вище, $t_s = 0, 1$ - напрям другої грані відносно першої (0 - поворот праворуч відносно напрямку першої грані, 1 - поворот ліворуч відносно напрямку першої грані). Комбінація декількох *L-елементів* здатна відтворити відрізок будь-якого напрямку з урахуванням дискретності ґратки.

Звичайно декілька однотипних *L-елементів*, що відрізняються тільки параметром l , розташовані компактно один за одним. Тому групу *L-елементів* доцільно описувати як $G = N, t, t_s \{l_n\}$, $n = 1, 2, \dots, N$. Відзначимо, що група *L-елементів* є складовою час-

тиною контуру, причому L-елементи, що її утворюють, мають однаковий напрям перших, "довгих"



Мал.1

граней. Тобто група L-елементів є одномірним об'єктом, у той час як контур є двомірним об'єктом.

ЕКСПЕРИМЕНТ

Для перевірки запропонованого методу створений пакет програм, який планується використати для обробки та дослідження зображень медичних препаратів, отриманих у системі Онкотест. На малюнках наведені приклади обробки зображень. На мал.1 показаний довільний бінарний малюнок при наявності поміх. На мал.2 результат роботи програми по кодуванню об'єктів малюнка, усуненню поміх згідно заданого заздалегідь параметру, що регламентує лінійний розмір поміхи, відновленню контуру. Об'єм масиву зображення відповідно 7062 та 4120 байтів.

ВИСНОВКИ

Вже перші кроки по застосуванню елементів довжини або L-елементів як елементів відрізків прямих, що утворюють контури, подають певні надії - досягнуто економію близько як у 2 рази порівняно з [1,2,3,4]) при кодуванні зображень. Безперечно, більша економія пам'яті порівняно з іншими є значною перевагою метода, що пропонується, особливо, коли мова йде про необхідність обробки та зберігання великої кількості зображень. Також як і економія часу при обробці та передачі зображень крізь канал зв'язку.

Але інші очікувані переваги видаються більш привабливими. По перше це те, що завдяки застосуван-

ню груп L-елементів виявилось, що двомірний об'єкт - контур природно розбивається на одномірні



Мал.2

частини - групи L-елементів. Таким чином є можливість двомірну по суті задачу звести до ряду одномірних.

Крім того після виключення поміх є можливість перейти до визначення контуру, як послідовності відрізків прямих та довільних кривих.

ЛІТЕРАТУРА

1. А.с. № 1354225 СССР. Устройство для считывания графической информации. / В.Г.Калмыков, Н.В.Рось, М.И.Шлезингер и др. Заявлено 10.01.1986. Зарегистрировано 22.07.1987.
2. А.с. № 1522256 СССР. Устройство для считывания графической информации. / В.Г.Калмыков, Н.В.Рось, М.И.Шлезингер и др. Заявлено 13.01.1987. Зарегистрировано 15.07.1989.
3. Мацелло В.В., Шлезингер М.И. Синтаксический анализ изображений в процессе построчного сканирования// Известия Академии наук СССР Техническая кибернетика,- Москва,1982, С 173-179.
4. Matsello V., Schlesinger M. On the Structural Analysis of Image Contours// Праці Третьої Всеукраїнської конференції "Оброблення сигналів і зображень та розпізнавання образів" - Київ, 1996, 26-30 листопада.
5. Батюк А., Балич Б., Вінтоняк Я., Цмоць І. Організація обчислювального процесу представлення і розпізнавання зображень в пороговому базисі Праці Третьої Всеукраїнської конференції "Оброблення сигналів і зображень та розпізнавання образів", Київ, 1996, 26-30 листопада.
6. Мартинес Ф. Синтез изображений. Принципы, аппаратное и программное обеспечение.- М.: Радио и связь, 1990.- 191 с.

ВИЯВЛЕННЯ ГРАНИЦЬ ЛЮДСЬКИХ ОБЛИЧ НА ЗОБРАЖЕННЯХ

Костянтин Нюнькін

Інститут проблем штучного інтелекту. 340048, Донецьк, вул. Артема, 118-б

Тел./Факс: (0622) 926 082. E-mail: kmn@iai.donetsk.ua

The first stage of boundary based methods of human faces recognition is the task of faces outlines detection in the images. The usual edge detectors give many false labels, and some edge parts can be lost. Therefore primary edges should be processed with the purpose of rise of their quality. The methods of such processing are the goal of the present research. The several algorithms are presented that allow to bind together the broken fragments as well as to remove the false connections of boundaries belonging to different scene objects.

ВСТУП

Розпізнання людини по зображенню його обличчя - цікава і важлива проблема, що привертає в останні роки увагу багатьох дослідників. Алгоритми аналізу людського обличчя є невід'ємною частиною систем ідентифікації осіб [1], систем стиску зображень людських облич, основаних на моделюванні голови [2], і систем візуального розпізнання мовлення [3].

Більшість алгоритмів розпізнання осіб припускають, що положення обличчя в кадрі відомо або ж воно легко виділяється на навколишньому фоні. Однак, в багатьох випадках це далеко не так. Тому задача виявлення та локалізації обличчя на зображенні є актуальною і має самостійне значення.

В цій роботі розглядаються методи попередньої обробки границь, що мають метою підготувати їх до наступного виявлення на зображенні людського обличчя.

1. ПОПЕРЕДНІ РОБОТИ

Багато алгоритмів локалізації на зображенні обличчя людини засновані на використанні границь. Найчастіше для пошуку границь використовується детектор Canny [4]. Знайдені границі потоншуються до товщини рівній одному пікселю. Після цього локалізуються точки перехрещень границь і виділяються сегменти що не перетинаються. Далі,

найбільш відповідні один до одного за напрямком сегменти зв'язуються у точках перехрещень [5].

2. ФОРМУВАННЯ ФРАГМЕНТІВ ГРАНИЦЬ

В результаті застосування до зображення серії алгоритмів, що забезпечують виявлення границь, контур обличчя, що міститься на зображенні, буде представлений в вигляді окремих фрагментів. Деталі обличчя, а також інші наявні на зображенні об'єкти також будуть представлені фрагментами своїх границь. Всі ці фрагменти можуть перетинатися або з'єднуватися один з іншим. Так, скажімо, контур голови може перейти в контур плеча, який, в свою чергу, може з'єднатися з контуром шафи або якихось інших об'єктів, що знаходяться за спиною людини. Крім того, контур обличчя може виявитися розірваним на окремі фрагменти (рис. 1). Звідси випливає необхідність попередньої обробки знайдених границь, що переслідує дві мети - розділити помилково з'єднані фрагменти границь різних об'єктів і, навпаки, з'єднати дрібні фрагменти границі одного об'єкту.

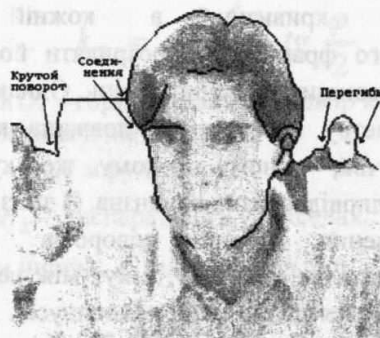


Рис. 1. Приклад границь

Така попередня обробка границь виконується в декілька етапів. Вхідними даними для неї служить карта границь, що являє собою "зображення", на якому виявлені точки границь відзначені білим кольором на чорному фоні. Лінії границь мають товщину рівну одному пікселю. Спочатку на карті

границь складаються точки пересічень і Т-подібних сполучень. Після цього відзначаються ті точки границь, що являють собою кінці фрагментів та видаляються поодинокі точки, які можуть заважати подальшим сполученням. Далі з'єднуються фрагменти, кінці яких відстоять друг від друга на один піксел. Після цього зв'язуються разом точки, що складають окремі фрагменти границь. Потім виконуються розриви фрагментів в точках різких поворотів (в цих точках скоріше за все з'єдналися фрагменти границь різних об'єктів). Для подальших перетворень фрагментів границь необхідна інформація про те, з якої сторони від фрагменту світла область, а з якої темна. Саме ця інформація визначається на черговому кроку обробки. Після цього, з урахуванням цієї інформації, можна виконати сполучення фрагментів, що являють собою частини однієї і тієї же границі. На заключним кроку попередньої обробки границь виконуються розриви фрагментів в точках перегинів. Сенс цієї операції в тому, що точки перегину можуть вказувати на перехід від однієї границі до іншої. Наприклад, перехід від контура голови до контура плеча (рис. 1). Розглянемо деякі з цих операцій детальніше.

Контури обличчя і голови не повинні мати різких поворотів. Точки різких поворотів, скоріше за все, є точками сполучення фрагментів границь різних об'єктів (см рис. 1). Мірою крутизни повороту кривої служить її кривизна:

$$k = \frac{x' \cdot y'' - x'' \cdot y'}{(x'^2 + y'^2)^{3/2}}, \quad (1)$$

де x' - перша і x'' - друга похідна по параметру. Параметром служить довжина кривої. Необхідно обчислювати кривизну в кожній точці розгляданого фрагменту і розривати його, коли кривизна перевищить заданий рівень. Однак, якщо в якості параметру виступає не довжина кривої, а кількість точок, одному і тому же куту між векторами відповідає різна кривизна. В зв'язку з цим, для визначення різких поворотів доцільно використовувати не кривизну, а кут між векторами, який визначається згідно теоремі косинусів.

На заключному кроці попередньої обробки границь виконуються розриви фрагментів в точках перегинів. Сенс цієї операції в тому, що точки перегину можуть вказувати на перехід від однієї границі до іншої. Наприклад, перехід від контура голови до контура плеча (рис. 1).

В кожній точці фрагменту обчислюється кривизна (1). Похідні по параметру, що входять в (1), обчислюються через кінцеві різниці. На ділянці

фрагменту, що має кривизну одного знаку, накопичується сума модулів кривизн окремих точок, яка лінійно зв'язана з довжиною дуги даної ділянки. Якщо ця сума на якійсь ділянці перевищує заданий поріг, ця ділянка являє собою окремий фрагмент.

Кривизна, що обчислена в сусідніх точках прямолінійної похилої границі, внаслідок дискретності зображення і границь може міняти знак, хоча насправді вона повинна зберігати нульове значення. Для того, щоб в подібних ситуаціях не реагувати на хибні зміни оцінки кривизни, вводиться поріг. Доки модуль оцінки кривизни менш цього порога, вважається, що кривизна дорівнює нулю.

ВИСНОВКИ

В цій роботі розглядаються методи попередньої обробки границь. Запропоновано декілька доповнень до існуючого алгоритму, що дозволяє зв'язати разом розірвані фрагменти контура обличчя і усунути хибні сполучення границь різних об'єктів сцени:

- Виконуються розриви границь в місцях різких поворотів, які скоріше за все є точками сполучення фрагментів границь різних об'єктів;
- При сполученні фрагментів, що являють собою частини однієї й тієї же границі, враховується інформація про те, з якої сторони від фрагментів знаходиться світла область, а з якої темна;
- Виконуються розриви фрагментів в точках перегинів, які вказують на перехід від одного контура до іншого.

ЛІТЕРАТУРА

1. Chellappa R, Wilson C. L., Sirohey S., Barnes C. S. *Human and Machine Recognition of Faces: A Survey.*/CS-TR-3339, 1994.-84p.
2. Strub L., Robinson J. *Automated Facial Conformation for Model-Based Videophone Coding.*//Proc. of IEEE Int. Conf. on Image Processing. - Washington D. C., 1995, v2.-p. 587-590.
3. Hennecke M, Prasad K., Stork D. *Using deformable templates to infer visual speech dynamics.*//Proceedings of the 28 Annual Asilomar Conference on Signals, Systems, and Computers, 1994.
4. Canny J. *A Computational Approach to Edge Detection.*//IEEE Trans. Pattern Analysis and Machine Intelligence, 1986, v. 8.-p. 679-689.
5. Sirohey S. A. *Human face segmentation and identification.*//Technical Report CS-TR-3176. - University of Maryland, 1993.-33p.

ПРО ВИМОГИ ДО СИСТЕМИ ПАСИВНОГО СТЕРЕОЗОРУ МОБІЛЬНОГО РОБОТА

Володимир Пчолкін

Інститут штучного інтелекту. 340048, Донецьк, вул. Артема, 118-б

Тел./Факс: (0622)926082. E-mail:vlad@iai.donetsk.ua

Detection of obstacles based on passive stereo vision is attractive for using in mobile robotic systems, but is very time-consuming in its implementation. Known algorithms of fast stereo matching apply some constraints to the stereo vision system. In the present paper the requirements to the geometric parameters of vision system are formulated, which both satisfy the mentioned constraints and provides desired accuracy of 3D scene detailization. The requirements to the depth map construction algorithm are also discussed.

ВСТУП

Система технічного зору є одною з найважливіших при створенні мобільного робота. В свою чергу, головним завданням такої системи є відтворення трьохвимірної сцени за даними, отриманими з одного чи кількох зображень. Найбільш очевидний шлях для вирішення цього завдання є використання стереозору. Як пасивна система, стереосистема дозволяє уникнути проблем, пов'язаних з інтерференцією випроміненого та відбитого проміння, що виникають при використанні активного зору. Разом з тим, при створенні зорової системи мобільного робота виникає низка специфічних проблем, пов'язаних з необхідністю розробки алгоритма, що, з одного боку, обробляє інформацію в реальному часі, а з іншого – забезпечує знаходження перешкод з необхідною точністю. Ці обставини накладають низку обмежень як на характеристики відеопідсистеми, так і на якість алгоритма. Деякі з цих обмежень обговорюються нижче.

1. ВИБІР СТЕРЕОБАЗИ

Більшість систем стереозору моделюються як пара однакових камер з оптичними вісями, розташованими в площині OXZ та паралельними OZ (див. рис. 1). Знаючи відстань між оптичними вісями обох камер (стереобазу), фокусну відстань і положення будь-якої точки на лівому та правому зображеннях, можливо знайти віддалення цієї точки від камер (глибину):

$$z = \frac{bf}{x'_n - x'_l} = \frac{bf}{\Delta x} \quad (1)$$

де b – стереобазу;

f – фокусна відстань;

Δx – різниця між положеннями точки на лівому та правому зображеннях вздовж OX (диспаратність).

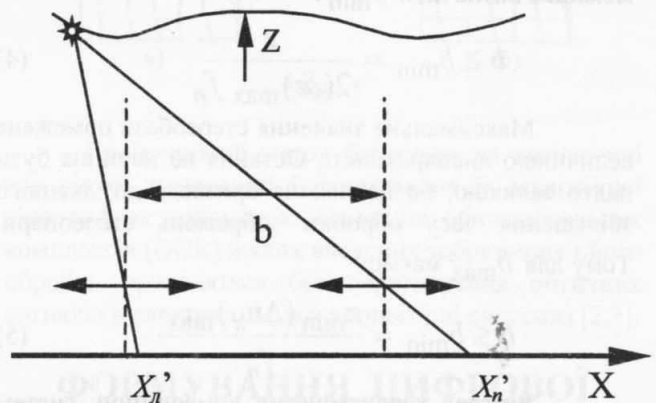


Рис. 1. Принципова схема розташування камер у типовій системі стереозору.

Під час обробки комп'ютерних зображень відстань вимірюється в пікселях, тому формулу (1) треба скорегувати, виразивши в пікселях диспаратність та фокусну відстань. Перевідний коефіцієнт k_x :

$\Delta x = k_x \Delta n_x$ дається співвідношенням:

$$k_x = \frac{2f}{\text{Max}X} \text{tg} \frac{\varphi}{2} \quad (2)$$

де $\text{Max}X$ – горизонтальний розмір комп'ютерного зображення в пікселях;

φ – кут зору камери;

Δn_x – диспаратність у пікселях.

Таким чином, z залежить від Δn_x :

$$z = \frac{bf n}{\Delta n_x} \quad (3)$$

де $f_n = \frac{\text{Max}X}{2 \text{tg} \frac{\varphi}{2}}$ – фокусна відстань в пікселях.

Для знаходження величини z треба знати b , f_n та Δn_x . Величина f_n є характеристикою

камери та пристрою відеоцифрування й визначається під час процедури калібрування. Калібрування камер необхідне також для застосування так званої епіполлярної моделі, що дозволяє звести двохвимірний пошук відповідних точок на зображеннях до одновимірного [1]. Величина ж стереобазис до деякої міри може задаватися вільно.

Мінімальне значення стереобазис b_{\min} обмежене вимогами точності визначення відстаней до віддалених об'єктів. Для відкаліброваних камер похибка в визначенні z пов'язана, головним чином, з помилками дискретизації зображень. Маючи $(\delta z)_{\max}$ - максимальну відносну похибку глибини, можливо визначити b_{\min} :

$$b \geq b_{\min} = \frac{z}{2(\delta z)_{\max} f_n} \quad (4)$$

Максимальне значення стереобазис обмежене величиною диспаратності. Остання не повинна бути надто великою, бо це інакше призведе до значного збільшення часу обробки зображень стереопари. Тому для b_{\max} маємо:

$$b \leq b_{\max} = \frac{z_{\min} (\Delta n_x)_{\max}}{f_n} \quad (5)$$

Розгляд характеристик комерційних систем стереозору [2] показує, що для забезпечення швидкості обробки на рівні 2 кадрів/с для зображень розміром 320x200 Δn_x не повинна перевищувати 64 пікселів.

Застосування формул (4) та (5) з урахуванням вищезазначеного визначає величину $b=0,1\text{м}$, як таку, що забезпечує точність 5% під час визначення глибини $z=10\text{м}$ і величину диспаратності ~ 30 пікселів для $z=1,5\text{м}$. Ця величина стереобазис близька до застосованої в [2].

2. ВИБІР АЛГОРИТМА

Головним завданням стереоалгоритма є відтворення трьохвимірної сцени за двома її проєкціями, фактично – визначення глибини для кожної точки зображень (побудова карт глибини). Головною вимогою, що висувається до такого алгоритма, поруч із надійністю, є велика швидкість обробки зображень, що важко забезпечити внаслідок великої обчислювальної складності. Всі відомі сьогодні системи стереозору, що працюють у режимі реального часу, використовують кореляційний підхід та спрощену епіполярну модель, згідно якій усі відповідні точки знаходяться на горизонтальних лініях та на однаковій висоті у обох зображень. Однак при цьому так чи інакше застосовуються апаратні особливості: суттєве використання ММХ-

технології [2], спеціальних процесорів та систем з паралельними обчисленнями [3,4].

Створення систем стереозору для мобільних роботів на основі IBM PC-сумісних комп'ютерів вимагає більш швидкого алгоритму ототожнення.

Прискорити пошук відповідних точок можливо, застосувавши, поруч з кореляційним підходом, багатомасштабний пошук, який дозволить суттєво скоротити область обчислення кореляції та розмір кореляційної матриці. Використання такого сумісного підходу було раніше з успіхом застосовано для розпізнавання людських облич [5].

ВИСНОВКИ

У цій праці проаналізовані вимоги до системи стереозору, придатної до застосування для обминання перешкод у мобільних робототехнічних пристроїв. Розглянуто як деякі технічні обмеження, так і вимоги до алгоритмів побудування карт трьохвимірних сцен.

- Запропоновані співвідношення, що пов'язують величину стереобазис системи пасивного стереозору з точністю визначення глибини, з одного боку, та величиною диспаратності (від якої однозначно залежить час пошуку відповідних точок на зображеннях стереопари), з іншого. Відмічено необхідність калібрування відеокамер такої системи.

- Зроблено оцінки величини стереобазис, що задовольняє обом співвідношенням. За оцінками отримано величину, близьку до використаної в комерційній системі стереозору.

- Запропоновано спосіб удосконалення існуючих стереоалгоритмів реального часу для більш швидкого ототожнення.

ЛІТЕРАТУРА

1. R.Y. Tsai. *An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision*//IEEE Computer Vision and Pattern Recognition.- 1987.- p.364-374.
2. <http://www.ptgrey.com/main.htm>
3. M. Bertozzi, A. Broggi. *GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection*//IEEE Trans. on Image Processing.- 1998.- V.7.- N.1.- p.62-81.
4. R. Klette, A. Koschan, K. Schluens, V. Rodehorst. *Surface reconstruction based on visual information*/Technical Report 95/6, Univ. of Western Australia. July 1995.
5. P.J. Burt. *Smart sensing within a pyramid machine vision*//Proc of IEEE.- 1998.- V.76.- N.8.- p.1006-1015.

ПОБУДОВА ЦИФРОВОЇ МОДЕЛІ МІСЦЕВОСТІ НА БАЗІ АНАЛІЗУ ТОПОГРАФІЧНИХ КАРТ ТА АЕРОФОТОЗНІМКІВ.

Михайло Семенюк

НВП "Струм", 286001, м.Вінниця, вул. Коцюбинського 11, кв.108.

E-mail: misha@post.vinnica.ua

В работе рассматривается метод преобразования информации, которая нанесена на топографических картах или фотоснимках полученных из космоса или самолета, в цифровую модель для последующей обработки на вычислительной технике. Метод основывается на принципе формирования залитых областей равной значимости на оптических вычислительных устройствах.

In the work the method is presented of transformation of the information, which was put on the topographic maps or photo pictures received from the space- or aircraft, into digital model for the consequent processing by the computers. The method is based on the principle of the formation of the filled regions of equal significance at the optical computerised devices.

ВСТУП

Інформація про місцевість використовується в якості вхідних даних в багатьох галузях господарства. При цьому рельєф є одним із основних геофізичних полів. Представлення його в вигляді цифрових матриць надає можливість обробки даних про нього за допомогою електронних обчислювальних машин. Представлення цих даних в цифровому вигляді набагато спрощує їх передачу по каналам зв'язку, подальшу обробку та зберігання.

Більшість алгоритмів формування цифрової моделі місцевості можна розділити на два основних метода:

- використання мануальних пристроїв перетворення аналогової інформації в цифрову шляхом реєстрації координат вузлових і допоміжних точок і, за допомогою численних методів отримання табличної (в координатах X,Y,H) чи аналітичної форми (функція висоти $H = F(x,y)$) представлення даних [1] (мал. 1а);

- оцифровування зображення шляхом построгового сканування робочого поля та реєстрація точок перетинання ліній сканування з ізолініями та лініями які обмежують ідентифікуючі ознаки об'єктів (в подальшому ця інформація може бути представлена як у вигляді масивів даних, так і в аналітичному (векторному) вигляді) [1] (мал. 1б).



Мал. 1

Пропонуємий метод базується на визначенні зон рівної значимості і формування їх ієрархічної зв'язності за допомогою оптичних обчислювальних комплексів (ООК) в яких введення зображення і його обробка проводяться без перетворення оптичних сигналів в електричні (око-процесорні системи) [2,3].

ФОРМУВАННЯ ЦИФРОВОЇ МОДЕЛІ РЕЛЬЄФУ

Основними об'єктами на топографічних картах і аерофотознімках є замкнуті геометричні фігури різної конфігурації - "поля". Межі "полів" представляють собою набір кривих та ламаних ліній, що математично можуть бути вираженими складними функціями двох змінних - x,y. Ці функції можна отримати або шляхом руху по межі об'єкту з реєстрацією координат [4], або шляхом неперервного сканування всієї площі зображення і подальшим ідентифікуванням належності окремих точок замкненим лініям. При цьому, найбільшу складність представляє собою ідентифікація точок, а отже і самої області.

Пропонуємий метод полягає в формуванні залитих областей на стадії вводу зображення карти в обчислювальний комплекс. Цей комплекс представляє собою структуру матричного типу (класу - багато даних, багато команд) в якому одночасно відбувається обробка масиву точок зображення, при цьому характеристика кожної точки зображення представляється функцією зв'язності із сусідніми точками:

$$C_{ij} = F(I_{n,m}), n=i-1,i,i+1; m=j-1,j,j+1, \quad (1)$$

де

I - вхідна інформація про аналізуєму точку.

Ця функція формується першим шаром обчислювальних елементів ООК. Крім того, на основі кодових позначень нанесених на топографічні

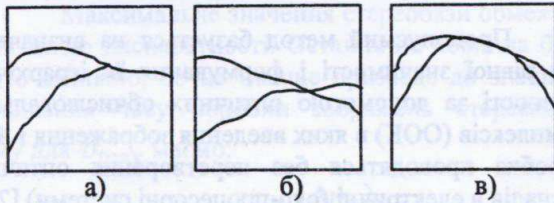
карти формується ознака напрямку висот (скату місцевості).

Другий шар ООК безпосередньо формує залиті області рівної значимості. Цей процес відбувається слідуочим чином. На підставі матриці зв'язності і кольору точки визначається належність точки до одного з елементів поля карти:

- межі (тобто, горизонталі топографічної карти, чи контуру зони фотознімка);
- внутрішній області.

При цьому, у зв'язку з обмеженою роздільною здатністю системи вводу, відбувається процес ідентифікації сполучень груп точок межі, які можуть бути віднесені до слідуочих класів:

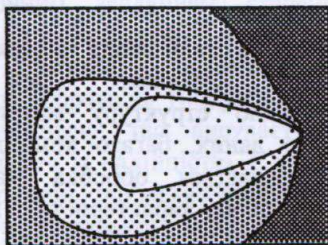
- Т-сполучення, примикання (розгалуження чи збіжність) однієї лінії до іншої (мал. 2а);
- Х-сполучення, примикання, або перетинання двох ліній (мал. 2б);
- співпадання ліній (мал. 2в).



Мал. 2

Далі відбувається процес зафарбовування областей зображення, що обмежені контурами (горизонталіями). Якщо при цьому "поле" знаходиться на краю зображення, то замикання лінії межі відбувається по прямій лінії яка співпадає з краєм зображення.

Перша точка верхнього лівого кута зображення яка належить "полю" приймається за нульовий рівень (в подальшому, при необхідності, він буде скоригований). Всі точки визначеного "поля" приймають значення кольору цієї точки. Точки сусіднього "поля" в залежності від знаку градієнту висоти приймають значення на одну дискрету більшу чи меншу. Такий метод дає однозначне визначення полів навіть при наявності випадків зображених на мал. 3.



Мал. 3

Після формування залитих областей – "полів" – проводиться корекція нульового рівня висоти і у відповідності до нього – значень всіх інших "полів". Для цього проводиться ідентифікація "нульового поля" (в ручному чи автоматичному режимі) і віднімання його значення від значення кожного поля.

В наступному шарі ООК відбувається безпосереднє формування цифрової моделі рельєфу (градієнтне зафарбовування полів). Висота кожної точки місцевості визначається за формулою:

$$H_{i,j} = H_k^z + \text{sign}(k) \cdot \Delta H_k^z \cdot \frac{\sqrt{(x_i - x_{k,i})^2 + (y_j - y_{k,j})^2}}{\sqrt{(x_{k+1,i} - x_{k,i})^2 + (y_{k+1,i} - y_{k,i})^2}}, \quad (2)$$

де

- H_k^z - значення висоти k-горизонталі;
- $\text{sign}(k)$ - напрямок скату висот;
- ΔH_k^z - різниця значень висот між горизонталіями k та k+1;
- $x_{k,i}, y_{k,j}$ - координати кінців радіального відрізка, що сполучає горизонталі.

За отриманими значеннями можливе як площинне моделювання розміщення об'єктів на місцевості завдяки рівномірному переходу кольору (висоти) від однієї точки до іншої, так і формування об'ємного зображення рельєфу.

Висновки

За результатами проведених досліджень можна зробити висновки, що використовуючи метод формування областей рівної значимості можливе оперативне формування моделі місцевості за топографічними картами та фотознімками зробленими з повітря та космосу у вигляді масивів висот рельєфу, графів лінійних об'єктів; проведення об'ємного моделювання при проведенні проектних робіт по прокладенню автомобільних та залізничних шляхів, ліній електропередач, трубопроводів, тощо; а також моделювання перебігу різних природних явищ в залежності від характеру місцевості. Крім того, даний метод дозволяє проводити архівацію та збереження зображень місцевості.

Література

1. Назаров А.С. "Методы создания цифровых карт и планов, и их использование в системах автоматизированного проектирования". Материалы НПК, Минск, 1986 г.
2. Кожем'яко В.П., Білан С.М., Семенюк М.В., Савалюк І.М. "Пристрій для виділення контуру багатоградаційного зображення". Патент № 95041856, 19995 р.
3. Кожем'яко В.П., Білан С.М., Семенюк М.В., Савалюк І.М. "Система розпізнавання образів на основі багатопарових матричних структур". Праці ІІ Міжнародної конференції УкрОБРАЗ-96, Київ, 1996 р.
4. Теренчук А.Т., Семенюк М.В., Савицкий А.В. "Алгоритм и устройство для определения параметров объекта сложной формы." Депонировано в ГНТБ Украины № 1761 Ук 93, 16.08.93

EXPERIMENTS ON SHAPE RECOGNITION USING MULTISCALE STRUCTURE MATCHING

Authors: Laura Davlea, Viorel Olariu
 Electronic Research Institute, Iasi, Romania
 Address: Lascar Catargi, Iasi, Romania
 E-Mail: ldavlea@mail.dntis.ro

Abstract: This paper addresses the problem of finding a common structure to many deformed samples of the same class. The algorithm used is the multiscale structure matching. We have experimented a new distance measure, based on the curvature of the contour.

I. INTRODUCTION

Learning visual models from samples of the same class is one of the central problems in pattern recognition and computer vision. This paper addresses the problem of generating a model from given samples, using a multiscale convex/concave structure of the shapes, and the extraction of the optimum scale structure common to shape samples of the class.

There are many previous works in which shapes are described by symbolic representations, which means those parts and parts relationships are in the form of attributed relational graphs.

The method described here does not need apriori knowledge about object shapes, reduces the computational complexity about graph models.

II. THEORETICAL FRAMEWORK

A shape contour can be expressed as two periodic functions $C(t)=(x(t),y(t))$, where x,y are the coordinates along the contour. An evolved version of this contour is defined using the convolution between $c(t)$ and the Gaussian kernel:

$$\begin{aligned} \Gamma_{\sigma} &= ((X(u,\sigma), Y(u,\sigma)), \quad u \in (0,1) \\ X(u,\sigma) &= x(u) \otimes g(u,\sigma); \\ Y(u,\sigma) &= y(u) \otimes g(u,\sigma); \end{aligned} \quad (1)$$

$g(u,\sigma)$ is the gaussian kernel.

The curvature along the contour is:

$$k(u) = \frac{x_u y_{uu} - x_{uu} y_u}{(x_u^2 + y_u^2)^{3/2}}$$

where u is the normalized length along the contour, and x_u, y_u, x_{uu}, y_{uu} , are the derivatives of the functions $x(t)$ and $y(t)$.

The points where $k(t,\sigma)=0$ are inflection points, and the segments between them are convex/concave structures.

The shapes can be represented as sequences of segments that have some characteristics. In our case these characteristics are the ratio l_i/L , and DKc_i , where

l_i = the length of the segment

L = total length of the contour

DKc_i = the variation of the curvature along the contour.

We consider two shapes A and B defined as two sequences of segments:

$$A^h = a_1^h, a_2^h, \dots, a_N^h$$

$$B^k = b_1^k, b_2^k, \dots, b_M^k$$

h and k are associated with two different scale factor σ_h, σ_k .

Increasing the value of σ , odd consecutive segments can be replaced by one segment. For the two shapes it can be two different scales for which there exist an optimum matching according to a distance measure defined between the two shapes.

The problem is to match segments of different shapes by finding corresponding segment pairs between them.

The matching procedure involves the definition of a dissimilarity measure between a_i , and b_j segments as:

$$d(a_i^h, b_j^k) = \frac{|DKc_i^h - DKc_j^k| \cdot ||i^h - j^k||}{|DKc_i^h| + |DKc_j^k| \cdot (||i^h| + |j^k|)} \quad (2)$$

where DKc_i^h is the signed difference between the maximum and the minimum value of curvature along the l_i segment.

We denote a sequence of $(2n+1)$ consecutive segments of a shape, $a(i-2n | i)$, and a sequence of $(2m+1)$ segments of shape B, $b(j-2m | j)$.

We consider that, $a(i-2n | i)$ at the finest scale can be replaced by a_i^h (segment i at scale h), and similarly $b(j-2m | j)$ with b_j^k (segment j of shape B, at scale k).

The multiscale segment dissimilarity between $a(i-2n | i)$ and $b(j-2m | j)$, can be formulated as:

$$D(a(i-2n | i), b(j-2m | j)) = d(a_i^h, b_j^k) + d_A(a(i-2n | i) \rightarrow a_i^h) + d_B(b(j-2m | j) \rightarrow b_j^k) \quad (3)$$

The first term is the one to one dissimilarity between segments a_i and b_j , and the others represent the cost of replacing $a(i-2n | i)$ with a_i and $b(j-2m | j)$ with b_j .

Given many samples of the same class of shapes we can find a common structure of segments.

We consider a shape A_i within N samples. This shape is partitioned into n_{ij} subsets (each subset contains an odd number of segments). Then matching shapes A_i to A_j is performed by finding a correspondence between the two sets.

$$P_{ij} = \{a_i, w^j | w=1, n_{ij}\}$$

P_{ij} is a partitioned segment family of sets.

$a \in P_{ij}$ consists of one or an odd number of segments of the finest scale on the shape contour. Similarly we can define P_{ji} as the family of subsets that best describe the matching of A_j by A_i . In the case of N samples of the same class, the matching is performed $N(N-1)/2$ times.

We consider the union of P_{ij} :

$$Q_i = \bigcup_{j=1}^N P_{ij}$$

Q_i is named a partially ordered set. The interpretation for shape generalization is equivalent to find the maximal elements in Q_i . The maximal element $q \in Q_i$ is an element for which there is no q' with $q \subset q' \in Q_i$.

III. EXPERIMENTATION

We have tested this method on deformed industrial shapes made by a small number of segments (<15). We have used the curvature for the dissimilarity measure, reducing the computational effort.

The minimization of the criterion function (2) is performed for a value of $\sigma=5-7$.

The algorithm for obtaining pairs of segments has 2 stage:

Firstly we match the segments using the dissimilarity measure formulated in (2).

Then we calculate the dissimilarity between the replaced segments (at the finest scale) and the corresponding segment at the scale σ .

In the second recognition stage the dissimilarity used is in the form:

$$d_A(a(i-2n | i)) = \omega(d(a(i-2n | i-2n+1)) + d(a(i-2n+1 | i-2n+2)), \dots, d(a(i-1 | i))) \quad (4)$$

where ω is a weighting factor.

A ω value of 0.6 reduces the effect of noisy contour.

In the case of models generation, given N samples of the same class, The matching procedure used are described in the Fig. 1.

The model generation is performed in many stages, and finally the class model $M12.N$ is find. The matching between samples uses the equation (2).

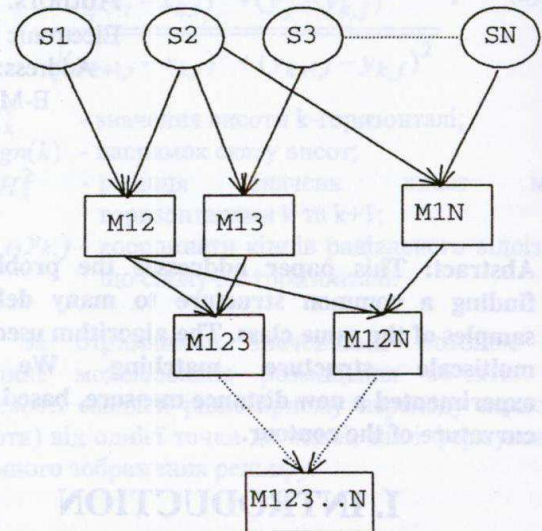


Fig. 1

IV. CONCLUSIONS

The method of shape recognition using a multiscale matching is suitable for deformed objects, and very robust with respect to position, orientation and scale change.

This method is computational less expansive than others such as those based on the curvature scale space which have the same objective.

V. REFERENCES

1. Robert Haralick and Linda Shapiro - *Computer and Robot Vision*
2. Milan Sonka and Vaclav Hlavac- *Image Processing, Analysis and Machine Vision*
3. Farzin Mokhtarian and Alan Mackworth - *A theory of multiscale, Curvature based Shape Representation for Planar curve* - IEEE Transactions on pattern analysis and machine intelligence, August 1992
4. Naonori Ueda and Satoshi Suzuki - *Learning visual models from shape contours using multiscale convex/concave structure matching* - IEEE Transactions on pattern analysis and machine intelligence, April 1993

Automatic matching of digital aerospace snapshots with using of Radon integrated transform

S.A. Stankevych

Kyiv Air Force institute

252186, Kyiv-186, Povitroflotsky avenue, 30

Phone: (044) 271-6236

Пропонується використання інтегрального перетворення Радона для будівництва бієкції сегментів комплексу цифрових зображень ділянки аероландшафту. Сподіване підвищення точності сполучення аерознімків має істотну роль при комплексному дешифруванні матеріалів видового аерокосмічного моніторингу.

The using of Radon integrated transform for construction of bijection between segments of a complete set of the digital images of an aerolandscape section is offered. The expected accuracy increasing of matching of aerosnapshots has an essential role for complex interpretation of imagery aerospace monitoring materials.

The matching of various digital images is the one important class of problems of automated processing of materials of imagery aerospace monitoring. These images are an aerospace snapshots obtained in a miscellany time, into various spectral zones, from systems with distinguish geometry of images creation or generally of various applicability – for example, during outlines binding of an aerosnapshot to a digital terrain map. The obtaining of synthesized digital images on materials of multizonal aerospace survey is impossible without high-precision preliminary matching of source zonal aerospace snapshots [4,5].

In existing systems of matching of digital images is applied morphing on a base points system. These points are determinates by the operator manually. In [3] the morphing system is represented as two components: the generalized linear transform, that nsure a main part of transformation, and residual non-linear elastic deformation by a field of dismatching vectors. A nontrivial problem is the automatic construction of a base points system of matched images. Usually for these purposes the chosen outlines of segments of digital images are used [1]. Defect of the mentioned approach is the rather low energy of an outline of a segment comparable with a noise level of an actual digital image. Traditionally, the reduction of high-frequency spatial noise of digital images is executed by the integrating operators.

One of such convenient operators represents an integrated Radon transform – is a generalized expansion

of a $f(x_n)$ function of initial n -dimension space $x_n \in X_n$ to its integrals along every possible hyperplanes ξ_{n-1} [6]:

$$F(\xi_{n-1}) = \int_{\xi_{n-1} \in X_n} f(x_n) dx_n \quad (1)$$

There is exist an inverse transform, that regenerating initial function by value of it Radon transform:

$$f(x_n) = \frac{2 \nabla_n^2}{(-4\pi)^{(n-1)/2} \Gamma(n/2)} \int_{x_n \in \xi_{n-1}} F(\xi_{n-1}) d\xi_{n-1} \quad (2)$$

where

$$\nabla_n^2 = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} \quad (3)$$

is the Laplace operator.

The Fig.1 illustrates a consecutive using of the direct and inverse Radon transform to a test digital image.

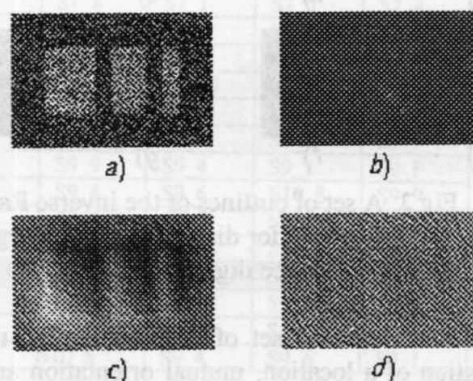


Fig.1. Integrated Radon transform:
a – the source image, b – the inverse transform (2), c, d – the consecutive using of the Hamilton operator ∇_n

The common defect of the competing integrated operators, for example – a median filtration, is the localization of using acts, that results to a diffusion of the boundaries of segments after derivation. It is illustrated a Fig.2 (c and d against g and h).

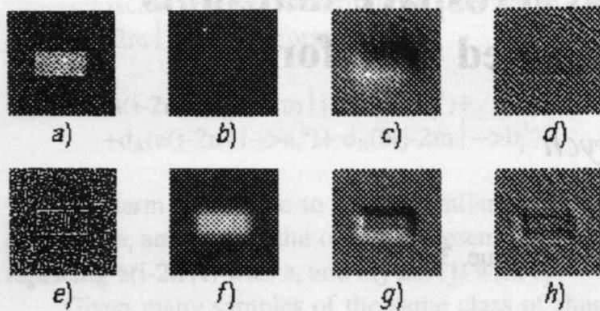


Fig.2. A comparison of Radon transform (b) with local filtration (f) of a source digital image (a)

It is obviously, that the recovery of the boundaries of a segment till a Fig.2d is possible with higher accuracy, rather than till a Fig.2h. An important feature of Radon transform using to processing digital images – is the capability, on the one hand, acts to limit only of some lines of initial plane and, accordingly, to allocate the boundaries of segments only collinear by it, and on the other hand – to use as a hyperplane, along which the integration conduct, any arbitrary points set of an source image, and not just lines set. The given circumstance allows to apply such generalized Radon transform to allocate segments of the certain beforehand specific configuration on a digital image.

The inverse Radon transform forms on a digital image some regular texture including a set of outlines, appropriate to segments configuration, that the Fig.3 is illustrated.

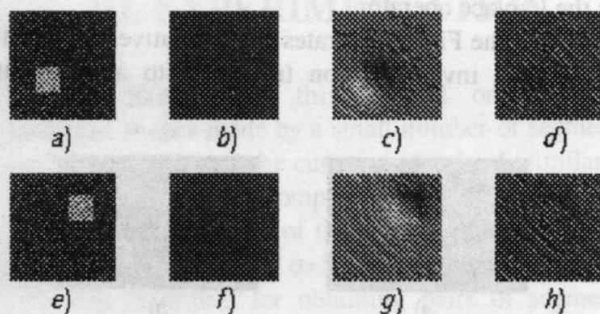


Fig.3. A set of outlines of the inverse Radon transform for displacement of a segment of an source digital image (a) - (e)

The indicated set of outlines can be used for clarification of a location, mutual orientation and scale of images of segments. The given circumstance can be useful for the stability increasing of construction of bijection between the matched images. And the correctly constructed bijection is the guarantee of precision matching of interpreted aerospace snapshots [2].

Thus, the sequence of direct and inverse integrated Radon transform provides a high-frequency spatial filtration of digital images, saving during of it the precise boundaries of segments and forming an additional set of signs of their shape and orientation.

With reference to aerospace snapshots the described property can represent itself as a base for construction of bijection between different images of one section of an aerolandscape by the analysis of topological invariants of segments [1]. The order of operations for automatic matching of digital aerospace snapshots is explained to a Fig.4.

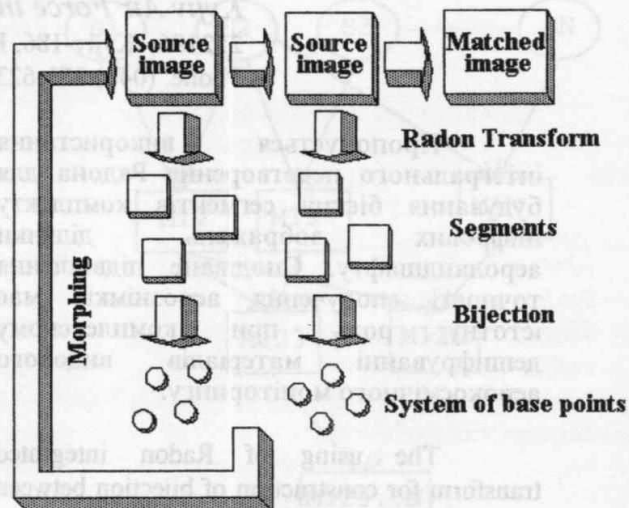


Fig.4. The scheme of automatic matching of digital aerospace snapshots

REFERENCES

1. Станкевич С.А. Модели автоматического сполучення сегментованих дискретних зображень // Праці другої Всеукраїнської міжнародної конференції з обробки сигналів і зображень та розпізнавання образів.- Київ: ІК НАН України, 1994.- С.167-169.
2. Станкевич С.А. Оценка качества синтезированных цифровых изображений // Информационные технологии в дешифрировании изображений: Классификация и оценка эффективности.- Киев: МО Украины, 1995.- С.36-51.
3. Станкевич С.А., Шестаков К.В. Модель биективного совмещения цифровых изображений // Проблемы повышения эффективности энергетических, разведывательных и управляющих комплексов / Материалы военно-научной конференции.- Киев: МО Украины, КИ ВВС, 1996.- С.53-56.
4. Stankevich S.A. The Linear Models of Optimal Synthesis of the Discretic Zonal Images // Proceedings of the Third All-Ukrainian Conference by Signal/Image Processing and Pattern Recognition.- Kyjiv: NAS IC, 1996.- P.149-152.
5. Станкевич С.А. Оптимальное факторное синтезирование зональных цифровых изображений // Проблеми здобування, збору та обробки даних повітряної розвідки / Збірник наукових праць.- Київ: МО України, КІ ВПС, 1998.- С.18-22.
6. Хелгасон С. Преобразование Радона: Пер. с англ.- М.: Мир, 1983.- 152 с.

Geometric Features for Face Recognition

V. Starovoitov and D. Samal

Institute of Engineering Cybernetics, Surganova 6, Minsk, 220012, Belarus

{valeryS,samal}@newman.basnet.minsk.by

A study of geometric approach to face recognition problem is presented. The Euclidian distance was used as a similarity measure in the feature space. The recognition rate equals 90% on the database of 50 images.

INTRODUCTION

There is a lot of approaches to face recognition problem. We have analysed most of them in the survey [1]. The oldest one is based on the face feature points, which are also called the anthropometric points. There are some merits and demerits of this approach. It is need about 40 anthropometric points to describe a face quite robustly. These points are known and used by criminalists widely. The problem is to find the optimal geometric feature set for a face description. It means the images of similar people or images of one person must be closer to each other in the feature space than to images of other persons. The difficulties are in great quantities of geometric feature variations. For example, there are 780 of all possible segments what may be drawn between 40 points. Beside that, geometric features may be presented by perimeters and areas of some figures formed by the points and so on. It makes choosing difficult. The aim of our study is to obtain several steady features sets, to compare them and to find the best one. This paper presents the first steps of our study.

We can present faces as patterns in the feature space. That is why the simplest measure in the recognition process may be the Euclidian distance between the points. To check efficiency of this idea we have provided several experiments. We used the ORL face database. The image name - SI_J means the J-th image of I-th person.

Table 1. Points and Distances used

N	Distance	N	Distance
1	$((ER2,ER1)+(EL1,EL2))/2$	16	$((ER2,ER1)+(EL1,EL2))/2$
2	$((ER4,ER6)+(EL4,EL6))/2$	17	$((LL2,FL11)+(LR2,FR1))/2$
3	$(EL5,ER5)$	18	$((ER2,NR1)+(EL2,NL1))/2$
4	$(NL2,NR2)$	19	$(BL3,BR3)$
5	$(LL2,LR2)$	20	$((ER1,BR1)+(EL1,BL1))/2$
6	$(FL1,FR1)$	21	$((EL2,BL3)+(ER2,BR3))/2$
7	$(LL1,FL1)$	22	$((ER2,ER1)+(EL1,EL2))/2$
8	$(LR1,FR1)$	23	$((ER6,BR4)+(EL6,BL4))/2$
9	$(LL1,LR1)$	24	$((BR5,BR4)+(BL5,BL4))/2$
10	(LM,LU)	25	$((BL5,BL1)+(BR5,BR1))/2$
11	(LM,LD)	26	$((BL5,BL1)+(BR5,BR3))/2$
12	(LD,TL)	27	$(EL2,NN)$
13	(NN,LU)	28	$(ER2,NN)$
14	$(NL1,NR1)$	29	$((ER2,LR1)+(EL2,LL1))/2$
15	$((EL5,NL2)+(ER5,NR2))/2$	30	$((EL1,EL5)+(ER1,ER5))/2$

EXPERIMENT 1

First we tested images of 5 people which were organized as 5 clusters. Every cluster included 5 photos of one person. Total number of images was 25. In the experiment the goal was "to find 5 nearest images". Carrying out the experiment, we measured 45 distances between 35 anthropometric points (Fig. 1) for every image. The distances were transformed into 30 features, Table 1. After that, the features were normalized by dividing on the distance between centers of the left and right irises. Describing the faces as points in 30-d feature space, we calculated the distances from every image to every other one and found 5 nearest images. The results are presented in Table 2.

Table 2. Results of experiment 1

Tested Image	Nearest to the tested image				
	1	2	3	4	5
S5_6	S5_9	S5_2	S5_8	S9_4	S9_8
S5_10	S7_5	S7_4	S7_1	S6_5	S6_4
S5_9	S5_6	S5_8	S5_2	S10_5	S9_4
S5_2	S5_6	S5_9	S5_10	S7_5	S9_8
S5_8	S5_9	S5_6	S5_2	S10_5	S10_6
S6_5	S6_4	S6_7	S6_6	S10_10	S10_8
S6_3	S6_6	S6_5	S6_4	S6_7	S10_10
S6_4	S6_5	S6_6	S6_7	S9_8	S9_9
S6_6	S6_7	S6_3	S6_5	S6_4	S10_10
S6_7	S6_5	S6_6	S6_4	S6_3	S10_10
S7_5	S7_1	S7_4	S5_10	S7_8	S7_7
S7_7	S7_8	S7_1	S7_5	S7_4	S5_10
S7_8	S7_7	S7_1	S7_5	S7_4	S5_10
S7_1	S7_5	S7_8	S5_10	S7_4	S7_7
S7_4	S7_5	S5_10	S7_1	S7_8	S7_7
S9_9	S9_8	S9_4	S9_6	S9_1	S10_5
S9_8	S9_9	S9_4	S9_6	S9_1	S6_4
S9_6	S9_9	S9_4	S9_1	S9_8	S10_1
S9_1	S9_4	S9_6	S10_8	S9_9	S10_5
S9_4	S9_1	S9_9	S9_6	S9_8	S10_8
S10_5	S10_6	S10_8	S9_4	S9_1	S9_9
S10_10	S6_5	S6_6	S6_4	S6_7	S6_3
S10_6	S10_5	S10_8	S9_1	S9_4	S9_6
S10_8	S9_1	S10_5	S10_1	S10_6	S9_4
S10_1	S10_8	S9_4	S9_6	S9_1	S10_5

There are two images S5_10 and S10_10 for which 5 nearest to them are not images of the same person. (see Table 2.). There is just one case (S10_8) when the nearest face is belonged to other person. The recognition accuracy was $23/25 = 92\%$. By the word, the people whose images were used in first part of experiments were not similar to each other.