

VI

Оброблення та розпізнавання рукописної, креслярської та графічної інформації. Аналіз фотозображень і сцен

Hand-Written, Drawing and Graphic Picture Information Processing and Recognition. Photopictures and Scene Analysis

One Generalization of Context-Free Grammars and Its Application to Structural Analysis of Drawings

IVAN AKSAK, VOLODYMYR KYJKO, VJACHESLAV MATSELLO, MICHAEL SCHLESINGER

NAS Institute of Cybernetics

40 prospekt Akademika Hlushkova, 252022 Kyiv

Tel.: (044) 266-2569 Fax: (044) 266-1570

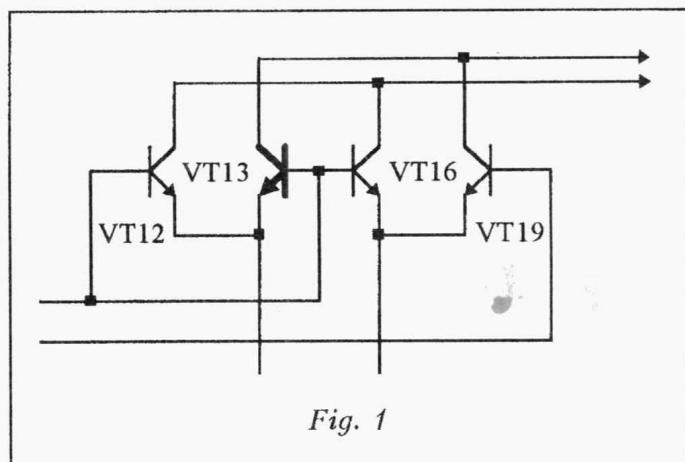
schles%image.kiev.ua@ts.kiev.ua

Ivan Aksak, Volodymyr Kyjko, Vjacheslav Matsello, Michail Schlesinger. One Generalization of Context-Free Grammars and Its Application to Structural Analysis of Drawings.

A problem of parsing of image elements set is considered. A new algorithm of hierarchical description of the set of image elements is proposed, and the result of its testing on real handdrawn pictures is adduced. The new algorithm is based on the certain generalization of context-free grammars, that has greater generating possibility than ordinary context-free grammars.

Description of applied problem. The proposed algorithm is intended for making the descriptions of drawings. Saying "description" we mean, that it is necessary to define that, for example, Fig.1 consists of: the first object, which consists of transistor's symbol and inscription, inscription consists of transistor's conventional name and number, transistor's conventional name consists of letters **V** and **T**, number consists of figures **1** and **2**; the second object, which consists of ..., etc.

To produce such description, some set of mergings must be fulfilled: at first the letters **V** and **T** are to be united into the word **VT** (transistor's conventional name) and figures **1** and **2** are to be united into the number **12**. Then the word **VT** and number **12** are united into inscription, and at last the graphical symbol of transistor is united with the inscription as the final compound object, which has its own certain meaning in the image, and cannot be united with any object. An obvious approach when elements are united if: 1) they can be united in general (for example **V** with **T**, but not **V** with **—**), 2) they are nearly located in the image, is not correct by more detailed consideration, and can solve the problem only in some special



cases. The reason of that consists in that although the whole image is quite unambiguous, its local parts are usually ambiguous. For example, although inscription **VT13** (Fig. 1) is near to two transistor's symbols, it is evident, that it relates to transistor drawn bold, because in other case it would be impossible to relate inscription **VT12** to any other transistor. It is easy for a human to parse such image, but to do it automatically the speculations of such type must be formalized.

We consider the problem as a specific problem of syntactic analysis of objects, where the set of allowable objects is defined by certain type of attribute grammar. The object under parsing is the set of primary image elements, which are obtained using some procedure of image elements recognition, for example [1]. So the image in Fig. 1 is supposed to be represented as a list, which consists of all transistor's symbols, all figures and all letters, which exist in the picture.

The formal statement of the problem. Main definitions. Let \mathbf{N} be a finite alphabet of names, $\mathbf{N}^P \subset \mathbf{N}$ be a subset of primary names and $\mathbf{N}^F \subset \mathbf{N}$ be a subset of final names. Let \mathbf{A} be a linear space, in which the subset \mathbf{A}^F is defined. We consider an object as a pair (n, a) , where $n \in \mathbf{N}$, $a \in \mathbf{A}$. The pair (n, a) is called an identifier of the object. The element n of the identifier is a name, and element a is an attribute of the object. Objects, whose names $n \in \mathbf{N}^P$ are called primary, other objects we will call compound. Each compound object consists of some nonempty set of primary objects.

We will distinguish allowable and not allowable objects using the following formal construction, that is the certain generalization of context-free (CF) grammars by Chomsky [2].

Let $\mathbf{P}^3 \subset (\mathbf{N} \times \mathbf{A}) \times (\mathbf{N} \times \mathbf{A}) \times (\mathbf{N} \times \mathbf{A})$ and $\mathbf{P}^2 \subset (\mathbf{N} \times \mathbf{A}) \times (\mathbf{N} \times \mathbf{A})$ be two relations, that define an identifier to be referred to compound object, i.e. to the set of primary objects.

Definition 1. A set \mathbf{O} of primary objects, i.e. compound object, can receive the identifier (n, a) in the system of relations \mathbf{P}^3 and \mathbf{P}^2 , if one of the next two conditions is true:

1. The compound object \mathbf{O} consists of the single primary object (n, a) .

2. There exists such compound object \mathbf{O}' with the identifier (n', a') , that $((n, a), (n', a')) \in \mathbf{P}^2$.

3. There exist such two compound objects \mathbf{O}' and \mathbf{O}'' with the identifiers (n', a') and (n'', a'') , that: a) $\mathbf{O}' \cap \mathbf{O}'' = \emptyset$; b) $\mathbf{O}' \cup \mathbf{O}'' = \mathbf{O}$; c) $((n, a), (n', a'), (n'', a'')) \in \mathbf{P}^3$.

Let's define an allowability of compound object in the system of relations \mathbf{P}^3 and \mathbf{P}^2 .

Definition 2. A compound object \mathbf{O} , i.e. a set of primary objects, is allowable in the system of relations \mathbf{N} , \mathbf{N}^P , \mathbf{N}^F , \mathbf{A} , \mathbf{A}^F , \mathbf{P}^2 , \mathbf{P}^3 if the object \mathbf{O} can receive an identifier $(n, a) \in (\mathbf{N}^F, \mathbf{A}^F)$.

Definition 3. A set \mathbf{O} of primary objects is an allowable image, if such subsets $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_n$ exist, that $\mathbf{O}_1 \cup \mathbf{O}_2 \cup \dots \cup \mathbf{O}_n = \mathbf{O}$, $\mathbf{O}_i \cap \mathbf{O}_j = \emptyset$ ($i \neq j$) and each \mathbf{O}_i is an allowable compound object.

Definition 4. An image description is a set of derivation trees of the following type (Fig. 2):

1) leaves (L_1, L_2, L_3, L_4) of the tree correspond to primary objects. Each leave has one output. It means that corresponding primary object is contained in the object connected to the output;

2) root (**R**) of the tree corresponds to allowable compound object. Each root has one or two inputs. Corresponding allowable compound object consists of the objects connected to inputs;

3) each node (N_1, N_2, N_3, N_4) corresponds to allowable compound object. Each node has one output and one or two inputs. Corresponding compound object is contained in another compound object, connected to the output, and consists of the objects connected to his inputs.

To solve the applied problem described above, we use the proposed construction in the following way. The set of attributes \mathbf{A} is the set of 4-tuples of numbers, i.e. each object is represented by 5-tuple (n, l, u, r, d) . The first element n is a name. Numbers l, u, r, d are coordinates of the left, up, right and down sides of the minimal rectangle, which envelopes the corresponding object in the image.

The relation \mathbf{P}^3 is defined as the set of rules of type: $n_i = n'_i n''_i \Psi_i$, where i is a number of the rule, n_i, n'_i, n''_i are the names, Ψ_i is a relation defining the set of allowable 8-tuples of numbers $\{l', u', r', d', l'', u'', r'', d''\}$ for which this rule can be used. The relation Ψ_i has a form of a system of linear inequalities, that relates the elements of this 8-tuple. Attributes of the object in the left side of substitution are determined as the following:

$$l = \min(l', l''), \quad u = \min(u', u''), \quad r = \max(r', r''), \quad d = \max(d', d'').$$

The relation \mathbf{P}^2 is defined as the set of rules of type: $n_j = n'_j \Psi_j$, where j is a number of rule, n_j, n'_j are the names, Ψ_j is a relation defining the set of allowable 4-tuples $\{l', u', r', d'\}$ for which

this rule can be used. The relation Ψ_j also has a form of a system of linear inequalities, which relate the elements of this 4-tuple. Attributes of the object in the left side of substitution are the same as ones in the right side, i.e. $l = l'$, $u = u'$, $r = r'$, $d = d'$.

Statement of the problem. Let image be represented as the set of primary objects and the system of sets \mathbf{N} , \mathbf{N}^P , \mathbf{N}^F , \mathbf{A} , \mathbf{A}^F , \mathbf{P}^3 , \mathbf{P}^2 is defined. The task is to make a conclusion about allowability of the image, and if it occurs to be allowable, then make a description of the image.

This problem is difficult.

We propose a plausible algorithm, that finds an image description if this description is unique. The algorithm generates a description of the input image automatically, if this image is allowable and has only one possible description. The algorithm calls for human operator, when during the analysis it finds that the input image is not allowable, or that the input image can't be unambiguously described (has more than one possible description).

Brief description of the algorithm. At first a *binding up* procedure is executed over the list of image objects. Initially this list contains only primary objects. While operating the new compound ones are added to the list. All objects from the list are looked over. If according to relation \mathbf{R}^3 it is possible to create a compound object from the pair of objects and these objects do not embrace common objects, then the new compound object is created and added to the list. The objects from this pair will be considered as initial objects for new compound object, and from the other side, new object will be considered as a derivative for each object of the pair. If according to relation \mathbf{R}^2 it is possible to create a compound object from some existing object, then the new compound object is created and added to a list. By analogy, the object from which the new one was created will be considered as an initial object for the new compound object, and new object will be considered as a derivative. Procedure is finished when no new object can be created.

Then the *obliterating* procedure is applied to the list formed by the binding up procedure. All objects of the list except primary and allowable compound objects are looked over. The objects which have no derivatives are obliterated. When some object was obliterated, its initial objects no more have it as a derivative. The procedure is finished when no object can be obliterated.

After that all primary objects of the list are inspected. If primary object has no derivative, it means that the image is not allowable. In this case the message about the necessity of correction is sent to an operator. After correction the algorithm starts from the very beginning.

Then the procedure of *deleting* [3] is executed. The primary objects contained in the single allowable compound object are looked for. If such primary object was found, then we consider corresponding allowable compound object to be correctly recognized. It allows to delete all other allowable compound objects containing the primary objects, which are also contained in correctly recognized allowable compound object. As a result some more primary objects, contained in a single allowable compound object may appear. There are two possible consequences of image nonallowability in this step of the algorithm. The first is that primary objects appear which are not contained in any allowable compound object. In this case the message about the necessity of correction is sent to an operator. After correction the algorithm starts from the very beginning. The second case: it is impossible to delete any allowable compound object, but some primary objects contained in more than one allowable compound objects remain. It means that more than one allowable image description exists. In this case the message about the necessity of choosing a single description among all is sent to the operator.

If neither the first nor the second case occurs it means that the image is allowable in the grammar, and the list of objects is the description of the image, where the root of each tree is allowable compound object and the leaves are primary objects.

Concluding remarks. We have tested the algorithm on real handdrawn pictures. One of them is shown on Fig. 3.

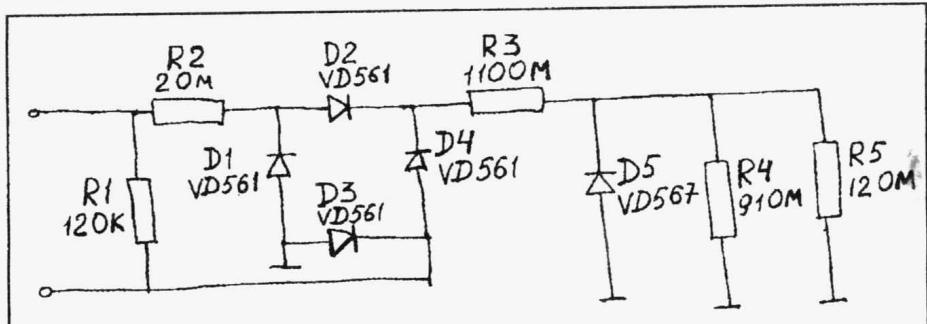


Fig. 3. This picture was scanned, preprocessed and recognized using [1]. Finally the procedure of parsing was applied to the results of recognition. The parsing procedure ran fully automatically (there were no calls to operator), and took 660 msec on Pentium 90 system.

References

1. V.M.Kijko. Recognition of Objects in Images of Paper Based Line Drawings // Proc. Of 3d Int. Conf. On Document Analysis and Recognition. - Montreal, 1995, Vol. 11, pp. 970-973.
2. Alfred V. Aho, Jeffrey D. Ullman. The Theory of Parsing, Translation and Compiling. - Prentice-Hall, Inc., Englewood Cliffs, N.J., Vol. 1, 1972.
3. M.I. Schlesinger. Formale Zugaenge fuer die Strukturanalyse von Grafischen Bildern // Musterkennung 1992, 14.DAGM-Symp. - Dresden, 14-16, September 1992, pp.10-19.



Метод оцінки якості формування дискретних фотографічних зображень

ІГОР БІЛЕЦЬКИЙ

Інститут військово-повітряних сил

252186 Київ, просп. Повітрофлотський 30
Тел.: (044) 271-6236

Ihor Bilets'kyj. The Method of an Estimation of Quality of Formation of the Discrete Photographic Images.

Method of an estimation of quality of formation of the discrete photographic images is offered. The essence of a method is a determining the recognition probability of a remote sensing object of aerolandscape by known distance up to it. The connection between distance up to remote sensing object of an aerolandscape and its recognition probability is provided by means of an instant vision field of system of remote sensing of an aerolandscape.

Для розв'язку задачі синтезу систем дистанційного спостереження аероландшафту необхідно мати модель формування дискретного зображення аероландшафту. При побудові моделі формування дискретного зображення аероландшафту критерієм якості виберем ймовірність розпізнання об'єктів дистанційного спостереження аероландшафту. Вибір дискретної моделі формування зображень аероландшафту зумовлений тим, що дискретна модель є загальною моделлю роботи для всіх систем дистанційного спостереження аероландшафту [1]; зображення аероландшафту обробляються в дискретному наборі точок [2]; фотографічні зображення аероландшафту мають обмежені розміри, які залежать від розміру кадру.

Загальну модель формування зображення системою дистанційного спостереження аероландшафту можна подати у вигляді [5]:

$$L'(x, y, t) = \iiint_{-\infty}^{\infty} a(x-\zeta, y-\eta, t-\tau) * L(\zeta, \eta, \tau) d\zeta d\eta d\tau + Lw(x, y, t),$$

де $L'(x, y, t)$ - сигнал на виході системи дистанційного спостереження аероландшафту; $L(\zeta, \eta, \tau)$ - сигнал на вході системи дистанційного спостереження аероландшафту; $a(x, y, t)$ - функція розсіювання системи дистанційного спостереження аероландшафту; $Lw(x, y, t)$ - функція шуму системи дистанційного спостереження аероландшафту.

У роботі [3] А.Н.Живічіним та В.С.Соколовим отримана залежність ймовірності розпізнання P об'єкта дистанційного спостереження аероландшафту від роздільної здатності Rc :

$P = \exp \{-(K\phi / B_{ob} * Rc)^2\}$, де $K\phi$ - коефіцієнт розпізнання форми об'єкта дистанційного спостереження аероландшафту; B_{ob} - максимальний геометричний розмір зображення об'єкта дистанційного спостереження аероландшафту. Оскільки коефіцієнт $K\phi$ для кожного об'єкта дистанційного спостереження аероландшафту не визначено, тому для побудови залежності ймовірності розпізнання P від детальності d необхідно використовувати емпіричні дані.

Для оцінки можливості системи дистанційного спостереження аероландшафту з детальності d по усьому полю зображення у робочому діапазоні відстаней до об'єкта дистанційного спостереження аероландшафту $L \in [L_{min}, \dots, L_{max}]$, залежним від технічних характеристик системи дистанційного спостереження аероландшафту, необхідно знайти вектор параметрів детальності \vec{P}_d : $\vec{P}_d = <\vec{P}_{\gamma_{kp}}, L(L \in [L_{min}, \dots, L_{max}])>$, де $\vec{P}_{\gamma_{kp}}$ - вектор параметрів миттевого поля зору γ_{kp} елементарного інформаційного каналу [1] системи дистанційного спостереження аероландшафту на краю зображення. Миттєве поле зору фотографічної системи (ФС) дистанційного спостереження аероландшафту знаходиться з відношень [1]: у центрі зображення $\gamma_{\text{ц}} = 1/\{2 * R_{c, \text{ц}} * f^* [K_{ob}/(2 - K_{ob})]^{0.36} * K\sigma\}$, де $R_{c, \text{ц}}$ - роздільна здатність ФС у центрі зображення при абсолютному контрасті об'єкт-фон, f^* - фокусна відстань фотооб'єктива, K_{ob} - контраст об'єкт-фон, $K\sigma$ - коефіцієнт урахування погіршення роздільної здатності дією вібрацій та дефокусування; на краю зображення $\gamma_{kp} = 1/\{2 * R_{c, \text{ц}} * f^* [K_{ob}/(2 - K_{ob})]^{0.36} * K\sigma * \cos^2(\beta_T)\}$, де β_T - кут візуалізації зображення. Таким чином, вектори параметрів миттевого поля зору ФС

дистанційного спостереження аероландшафту у центрі $\tilde{\mathbf{P}}_{\gamma_{\text{ц}}}$ та на краю $\tilde{\mathbf{P}}_{\gamma_{kp}}$ зображень знаходять з відношень: $\tilde{\mathbf{P}}_{\gamma_{\text{ц}}} = \langle R_{c,k}, f', \text{Коб}, K\sigma^* \rangle$, $\tilde{\mathbf{P}}_{\gamma_{kp}} = \langle R_{c,k}, f', \text{Коб}, K\sigma^*, 2\beta \rangle$, де $R_{c,k}$ - роздільна здатність ФС при контрасті об'єкта-фону Коб , 2β - кут поля зору фотооб'єктива.

Пропонований метод оцінки якості формування дискретних фотографічних зображень ґрунтуються на тому, що миттєве поле зору γ зв'язане з важливими технічними показниками ефективності, якими є детальність d та відстань до об'єкту дистанційного спостереження аероландшафту L , простим співвідношенням: $d = \gamma^* L$ [1].

Використовуючи емпіричні залежності ймовірності P розпізнання об'єктів дистанційного спостереження аероландшафту від детальності d , а також аналітичні залежності, що зв'язують L , γ , можливо графічно побудувати залежності $L(d)$, $P(d)$ (див. рис. 1).

Метод оцінки якості ФС дистанційного спостереження аероландшафту дозволяє: оцінити здатності ФС з детальності d по всьому полю зображення у робочому діапазоні відстаней L , який залежить від технічних характеристик ФС; визначити мінімальну ймовірність розпізнання P_{L_3} вибраних об'єктів по полю зображення, більш якої відстань до об'єкту дистанційного спостереження аероландшафту буде менше L_3 ; проводити найскоршу оцінку здатності ФС дистанційного спостереження аероландшафту по виявленню об'єктів на заданій відстані L_3 .

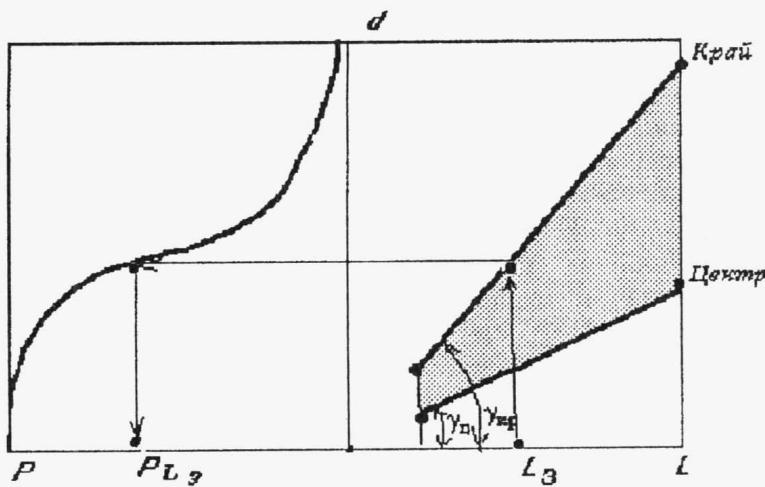


Рис. 1. Схема розрахунку якості фотографічних систем дистанційного спостереження аероландшафту

Література

1. Ребрин Ю.К. Оптико-электронное разведывательное оборудование летательных аппаратов. - Киев: КВВАИУ, 1988, 450 с.
2. Хорн Б.К.П. Зрение роботов: Пер. с англ. - Москва: Мир, 1989, 487 с.
3. Живичин А.Н., Соколов В.С. Дешифрирование фотографических изображений. - Москва: Недра, 1980, 253 с.
4. Гудмен Дж. Введение в Фурье-оптику. Пер. с англ. - Москва: Мир, 1970, 364 с.
5. Краснов М.Л. Интегральные уравнения. - Москва: Наука, 1975, 301 с.



Аналіз інформаційних технологій автоматичного розпізнавання і зведення об'єктів цифрової карти

ДМИТРО БЛИЗНЮК, МИХАЙЛО ПОПОВ

Інститут військово-повітряних сил

252186 Київ, просп. Повітрофлотський 30
Тел.: (044) 271-6560

Дмитрий Близнюк, Михаил Попов. Анализ информационных технологий автоматического распознавания и сводки объектов цифровой карты.

В настоящее время процесс создания цифровых карт достаточно длительный и трудоёмкий, поэтому новейшие разработки направлены на автоматическое распознавание образов и связей объектов и их структурирование. Применительно к задаче сводки смежных листов цифровой карты предлагается метод, объединяющий в себе метод автоматического распознавания объектов на карте по топологическим критериям и преобразование Хо.

Одним из джерел цифрових картографических данных є матеріали цифрового перетворення карт. На сьогоднішній день існує три методи оцифрування карт: шляхом планшетного дігітайзера (найбільш поширенний метод); методом послідовного відстеження ліній (напівавтоматичне оцифрування); з використанням сканера (найбільш перспективний).

Перетворювачі, засновані на двох перших методах, дають структуровані векторні дані. Оцифрування ручним способом - найбільш універсальний та дешевий спосіб, проте цей процес надто трудомісткий. Оцифрування методом послідовного відстеження ліній ефективне тільки для ізоліній. Недоліки цього способу - висока вартість та необхідність активної участі оператора. Автоматичне (сканерне) оцифрування дає велику економію часу, але оскільки на виході одержують растркові дані, то це потребує їх додаткового оброблення в інтерактивному режимі.

Оброблення сканерних даних у інтерактивному режимі досить трудомістке, тому новітні розробки спрямовані на автоматичне розпізнавання образів і зв'язків об'єктів та їх структурування. Стосовно картографії розпізнавання образів - це опис і класифікація літерно-цифрових ознак, умовних знаків, ліній, зокрема, меж об'єктів.

Автоматичне розпізнавання образів ґрунтуються на аналізі зв'язаних елементів матриць зображення за геометричними, топологічними та логічними критеріями та з наступним порівнянням одержаних параметрів з еталонними наборами параметрів. Незважаючи на прогрес в цій області існуючі програми розпізнавання картографічних образів ще недостатньо досконалі для ефективної автоматизації процесу утворення цифрової карти.

Розглянемо можливості застосування автоматизації для вирішення однієї з задач цифрового картографування - зведення суміжних аркушів цифрової карти на прикладах деяких з існуючих картографічних систем.

Система PC ARC/INFO американської фірми ESRI (Environment System Research Institute). Ця ГІС-оболонка використовується для розроблення окремих програмних продуктів, що потім об'єднуються у великі ГІС-системи вже на базі ARC/INFO для робочих станцій. Процедуру зведення суміжних листів цифрової карти можна представити наступною схемою: (1) одночасне візуальне зіставлення двох суміжних листів: візуальне порівняння листів до підгонки країв, підготовка покриття для підгонки країв (обрізання, розділення дуг на кутах тощо), виконання підгонки країв; (2) Перевірка значень параметрів у таблицях атрибутів об'єктів; (3) З'єднання покрить; (4) Усунення межі між листами карти.

На першому кроці візуально перевіряється: чи всі об'єкти мають пару на суміжному листі на протязі всієї межі і чи справді вони підходять один до одного. Це завдання вирішує оператор, використовуючи процедуру ARCPLOT. Після цього, в автоматичному режимі, виконується обрізка покрить для їх наступного об'єднання. Для цього застосовується операція топологічного перекриття CLIP. На цьому етапі відбувається призначення остаточних ідентифікаторів користувача, утворення нових міток і таблиці атрибутів полігонів (PAT-таблиці). Наступний етап - підгонка країв - виконується у інтерактивному режимі. Треба більш докладно зупинитися на цій процедурі, щоб поставити задачу автоматизації процесу зведення суміжних листів цифрової карти.

На перший погляд, використовуючи сучасні засоби оцифрування, які мають високу точність, можна було б очікувати на те, що підгонка країв не викличе великих труднощів, проте це не так. Помилки операторів при ручному оцифруванні (від зміщень до пропусків об'єктів), похибки на краях карти при скануванні, а також при поганій якості зображення, що пов'язано з властивостями картографічного паперу, призводять на практиці до того, що більшість об'єктів, що об'єднуються, зміщуються один відносно одного вздовж межі зв'язування. В PC ARC/INFO підгонка країв виконується з використанням процедури EDGEMATCH. Оператор проглядає узловж межі зв'язування усі вузли, що об'єднуються, встановлює "допустиме відхилення при підгонці" і задає опцію AUTOMATCH. Після завершення автозв'язування оператор знову проглядає всі вузли. Вузли, зміщення яких попадає у "допустиме відхилення", автоматично зв'язуються дугою, виділеною іншим кольором. До всіх інших вузлів, що не попали у "допустиме відхилення", застосовується опція MATCH (зв'язування кожної пари вузлів, що об'єднуються в інтерактивному режимі). Треба зазначити, що збільшення "допустимого відхилення" для досягнення автоматичного об'єднання всіх вузлів не призводить до позитивних результатів, бо зростає ймовірність зв'язування несуміжних вузлів, і тим більше, чим вище насичення об'єктами листів карти. Крім того, відзначимо, що якщо об'єднання покрить здійснюється не по прямій межі, то необхідно вручну вводити "тимчасові вузли" на кожному згині межі зв'язування. Остаточне об'єднання вузлів, що зв'язуються, виконується процедурою ADJUST, після завершення якої необхідно повторно перевірити правильність підгонки.

Крок 2 полягає у перевірці збігу визначень характеристик об'єктів, що зв'язуються, та у контролі узгодження атрибутів об'єктів на межі об'єднання листів карти. Для цього використовується процедура TABLES. Найбільш поширені помилки на цьому етапі: пропуск характеристик; змінена ширина поля характеристики; інший порядок характеристик.

Редактування PAT-таблиць здійснюється в інтерактивному режимі, що потребує значних витрат часу оператором. Кроки 3 і 4 процесу зведення виконуються автоматично, використовуючи процедури MAPJOIN і DISSOLVE .

Автоматична картографічна система “ПРІВАДА” розробки ІТК АН Беларусі. В принципі процедура зведення суміжних листів цифрової карти у системі “ПРІВАДА” аналогічна системі PC ARC/INFO, але існують деякі суттєві відміни, які можна визначити і як переваги, і як недоліки. До переваг можно віднести спроможність системи відображати тільки ті об'єкти, що виходять на межу листів карти і не попадають у “допустиме відхилення”, тим самим полегшуючи операторові виконання процедури зв'язування. До недоліків слід віднести таке: “допустиме відхилення” задається фіксованим для всіх вузлів, що зв'язуються, а вузли, що не попали у “допустиме відхилення”, зв'язуються оператором вручну шляхом “витягування мишею” додаткових дуг між суміжними вузлами; суміжні об'єкти, що розташовані з великим зміщенням, пов'язати описаним вище способом неможливо. Такі зміщення виправляються шляхом повернення листа карти на етапи редагування чи навіть повторного оцифрування цих об'єктів, на відміну від системи PC ARC/INFO, де подібні зміщення виправляються методом “гумового листа”. Цей метод дозволяє зсувати тільки ті дуги, що з'єднані з вузлами, що об'єднуються, всі ж внутрішні дуги залишаються без змін.

З раніше сказаного випливає, що ще значна частина робіт при виконанні зведення суміжних листів цифрової карти недостатньо автоматизована, особливо на етапі підгонки країв листів карти до об'єднання. Це необхідність постійного візуального контролю за відповідністю та точністю зв'язування кожної пари суміжних вузлів до і після операції автозв'язування, необхідність ручного зв'язування вузлів, які залишились після операції автозв'язування, редактування таблиць атрибутів об'єктів у інтерактивному режимі. Для автоматизації процесу зведення суміжних листів можна запропонувати такий підхід - об'єднання методу автоматичного розпізнавання об'єктів на карті за топологічними критеріями з перетворенням Хo. Суть методу полягає в тому, що використовуючи топологічні зв'язки, визначаються і ставляться у відповідність об'єкти вздовж межі листа карти, що підлягають зв'язуванню. Після цього доожної пари об'єктів, що зв'язуються, застосовується перетворення Хo. В результаті в площині параметрів Хo одержують два локальних максимуми, відповідні об'єктам що зв'язуються. Суміщуючи ці максимуми в площині параметрів Хo, із властивостей перетворення Хo одержують однозначне переміщення об'єктів на карті до їх об'єднання. Попередні розрахунки свідчать, що, використовуючи цей підхід, можна прискорити процес утворення цифрових карт.

Література

- Создание электронных карт: Обзорная информация // ЦНИИГАиК. - Москва, 1994, 60 с.
- Мартыненко А.И. Методы и средства автоматизации в картографии // Геодезия и картография, 1991, № 7, с. 44-45.
- Висмут А.С. Методика и компьютерная технология проектирования и составления электронных тематических карт // Изв. Вузов. Сер. Геодезия и аэрофотосъемка, 1992, № 1, с. 87-96.



Reading and Understanding Documents

L. DUFFY, H. EMPTOZ

INSA de Lyon,
Laboratoire RFV Bât 403

20 avenue Albert Einstein, 69621 Villeurbanne Cedex, France
Tél.: 72.43.83.83 poste 55.23 Fax : 72.43.80.97
email : duffy@rfv.insa-lyon.fr

L. Duffy, H. Emptoz. Reading and Understanding Documents.

In this summary, we suggest a new method of recognition and analysis of non structured documents, whose goal is the construction of a system for the automatic reading of documents, which feeds data base directly.

The problems bound to the document segmentation were studied by numerous teams of research [YU93] [WIESER93] [BELAID93] [ABLAMEYKO94], and the results produced are often very interesting and reliable.

As for the « intelligent analysis and understanding of documents » part, it remains limited to the world of called structured documents. This type of document obeys some rules of page layout enough precisely known and standardized, the logical and physical structure is already known. One could note in particular the works of Belaid [BELAID92] and of Ingold [INGOLD89].

As for non structured document they don't obey strong rules, and we possess little or no information on their logical and physical structure.

In order to raise this problem, we suggest a solution which, without being suitable for all the types of documents, could be applied to many of them.

Our survey essentially concerns the intelligent analysis and the understanding of document, we begin therefore with straightened, clean and correctly partitioned images.

First we chose to apply our solution to the synopses of magazines (Cf Fig 1), which presents the advantage, while being non structured documents, of being able to be expressed in terms of grammars.

Our system includes several levels :

① The first level involves the extraction of the typographic features of some blocks of print, excerpts of the image. We seek the grease, the style, the alignment of every block. The techniques employed are inspired by some of the work of Zramdini [ZRAMDINI95] and of his ApOFIS system. To every block, we therefore attach a list of attributes qualifying it, this allows us to create families of homogeneous blocks (Cf Fig 2).

② The second level consists of the labelling of these families, that is to give every family a semantic value (Cf Fig 3).

For do this we disengage of every type of documents, some rules regarding the logical structure and detects the ties between the different entities composing the document. We can thus construct a grammar characterizing the type of document with a hierarchy of the visual observed indications. All this information is placed in the information base corresponding to the type of document. The labelling is achieved by comparison of the visual indications excerpts and some constructed families, with the data contained in the information base. This base of information, common to all the documents of the same type, is constructed during a phase of training guided by any user.

In summary we could say that our logical interpretation of document is based on the physical and typographic structure of the document, thus that on the acquirements makes to the course of the establishment of the information base.

③ The third level of treatments concerns the validation of results. It necessitates both utilization OCR software and the inclusion of dictionaries adapted to the type of document.

Our work concentrates on the second level of treatment which must allow the logical part of a document, to be automatically labelled and construct, using a system of most ergonomic training possible, information bases corresponding to the types of documents that we want to recognize.

In summary we could say that our solution is interesting because it uses both some techniques which have already been proved, and some new methods emanating from artificial intelligence and some grammars. It allows for suppleness for the type of documents, thanks to the possibility that it offers also a creating an information base which is easily adapted to the document. What is more we are inclined to think that our system, with some modifications, will allow to suggest a solution for sorting documents.

--	--	--

Fig. 1. Brief binarized and partitionned

Fig. 2. Typographic segmentation

Fig. 3. Labelling the families of homogeneous blocks

Bibliography

1. [ABLAMEYKO94] ABLAMEYKO S. et OKUN O. Text séparation from Graphics based on compactness and area properties. *Machine Graphics ans Vision*, Vol 3, n°3, 1994, p. 531-541.
2. [BELAID92] BELAID A. et BELAID Y. *Reconnaissance de Formes*. Paris: InterEditions, 1992. 429 p.
3. [BELAID93] BELAID A. et AKINDELE O.T. A labelling approach for mixed document blocks. *International Conference in Document Analysis and Recognition*, Tokyo, Japon, 1993, Vol 4, p. 749-752.
4. [INGOLD89] INGOLD R. *Une nouvelle approche de la lecture optique intégrant la reconnaissance des structures de documents*. Thèse Doct. Sci. : Ecole Polytechnique de Lausanne, 1989, 130 p.
5. [WEISER93] WIESER J. et PINZ A. Layout and analysis : Finding text, titles, and photos in digitalimages of newspaper pages. *International Conference in Document Analysis and Recognition*, Tokyo, Japon, 1993, Vol 4, p. 774-777.
6. [YU93] YU C. L., TANG Y. Y. et SUEN C. Y. Document architecture language (DAL) Approach to document processing. *International Conference in Document Analysis and Recognition*, Tokyo, Japon, 1993, Vol 1, p. 103-106.
7. [ZRAMDINI95] ZRAMDINI A. et INGOLD R. «ApOFIS» : An a Priori Optical Font Identification System. *8th International Conference on Image Analysis and Processing*, Sanremo, Italie, 1995.



Два методи розпізнавання разом написаних слів

МИКОЛА КИРИЧЕНКО, ІГОР ГАЙДАЙЧУК, ВАЛЕРІЙ ФРАТАВЧАН

Державний університет

274012 Чернівці, вул. Коцюбинського 2
Тел.: (03722) 9-8448

Mykola Kyrychenko, Ihor Hajdajchuk, Valerij Fratavchan. Two Methods for Manuscripts Recognition.

Two methods for manuscripts recognition are presented. The first one is based on words representation as graphs and further grammatical analysis of obtained arcs. The second one works with words contours and codes contour by specially choosed elements.

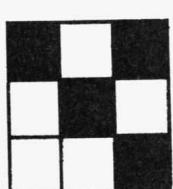
Розпізнавання разом написаних слів шляхом аналізу їх скелету. Алгоритм, що пропонується, можна розбити на три етапи: попереднє оброблення слова; перевірка гіпотез про наявність у слові символів з бази даних еталонів; ідентифікація слова за розпізнаними символами.

Перший етап включає в себе:

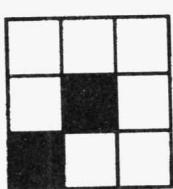
а) виділення слова із зображення тексту. Оскільки слово є разом написаним, то для знаходження його положення і розмірів зручно використовувати методи пошуку зв'язних областей [6];

б) проріджування слова (отримання його скелету). Серед сукупності методів, описаних у [5], найкращий за швидкодією і якістю результат дає ітераційна процедура, запропонована Накаші та Шингалом;

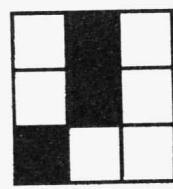
в) представлення скелету слова у вигляді графа, вузлам якого відповідають точки, що в околі з восьми сусідніх пікселів містять відмінну від двох кількість інших точок скелету (рис. 1); решта ж точок утворюють дуги.



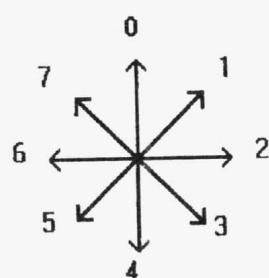
a



b



c



а) особлива точка - три сусіди;

б) вузлова точка - один сусід;

в) невузлова точка - два сусіди.

Rис. 1.

Rис. 2

Кожну дугу графа закодуємо напрямними відрізками Фрімана [1,5,7,8] (рис. 2): отримаємо для кожної дуги вектор напрямків її написання.

На другому етапі перевіряється наявність у слові символів, еталони яких знаходяться у заздалегідь навченій базі даних. Кожен еталон задається граматикою напрямків його написання, де напрямки, як раніше дуги скелета, також кодуються відрізками Фрімана. Розглянемо деякий фіксований еталон і для нього серед дуг графа виберемо ті, що задовільняють граматику еталона.

Нехай α - одна з таких дуг, e_α - її кінцева точка, P_α - стан автомата після закінчення розпізнавання дуги α . Серед дуг, що виходять з точки e_α , виділяємо ту, що відповідає граматиці еталона з початковим станом P_α . Отримуємо нову кінцеву точку $e_{\alpha'}$ і новий стан автомата $P_{\alpha'}$. Далі повторюємо цей процес знову, добавляючи на кожному кроці до послідовності проаналізованих дуг наступну таким чином, щоб нова дуга задовільняла граматику еталона. Якщо на деякому кроці буде досягнуто кінцевий стан граматики, то цим самим буде розпізнано у слові відповідний символ.

На третьому етапі залишилось вияснити, які з розпізнаних символів дійсно зустрічаються у слові, а які було отримано хибно. Ця проблема виникає у зв'язку з тим, що деякі символи є частинами інших: 'о' - 'а', 'а' - 'д', 'о' - 'в', 'п' - 'т' тощо. Крім того, написання разом різних символів також може привести до неоднозначності їх розпізнавання.

Рис. 3.

Як приклад, розглянемо слово 'тип':

Згідно з описаним вище алгоритмом, у ньому можуть бути ідентифіковані такі символи:

- один символ 'ш'

- один символ 'т'

- два символи 'п'

- один символ 'и'

- один символ 'г'

- два символи 'и'

Рис. 4.

В якості критерію для відновлення слова за розпізнаними символами авторами використано два припущення:

- ніякі два символи у слові не мають спільних дуг. Надалі вибірку з розпізнаних символів, кожна пара яких володіє вказаною властивістю, назовемо допустимою;
- символи повністю заповнюють граф скелету, за виключенням, можливо, рисок зв'язків між символами. Це означає, що та вибірка, яка утворює слово, буде мати найбільшу суму довжин своїх дуг серед усіх інших допустимих вибірок.

Для математичної постановки задачі введемо такі позначення:

a) c_i - сумарна довжина дуг i -ого символу;

b) $a_{ij} = \begin{cases} 1, & \text{якщо } i\text{-ий та } j\text{-ий символи мають спільні дуги,} \\ & 0 - \text{у протилежному випадку;} \end{cases}$

v) $x_i = \begin{cases} 1, & \text{якщо } i\text{-ий символ входить у слово,} \\ & 0 - \text{у протилежному випадку.} \end{cases}$

Тоді задача ідентифікації слова за розпізнаними символами запишеться у вигляді: знайти такий вектор $x = (x_1, x_2, \dots, x_n)$, де n - кількість розпізнаних символів, що є розв'язком задачі

$$\sum_{i=1}^n c_i x_i \rightarrow \max ; \quad \forall x_i = 1, x_j = 1: \quad a_{ij} = 0, \quad i, j = \overline{1, n} ; \quad x_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad i = \overline{1, n}. \quad (1)$$

Друга умова в (1) еквівалентна тому, що розглядаються лише допустимі вибірки, перша умова означає, що серед допустимих вибірок відшукається та, сума довжин дуг символів якої є максимальною.

Задачу (1-3) можна розв'язувати безпосередньо методом повного перебору, а можна звести її до еквівалентної задачі цілочисельного програмування:

$$\sum_{i=1}^n c_i x_i \rightarrow \max , \quad (2)$$

$$n \cdot x_i + \sum_{j \neq i} a_{ij} x_j \leq n \quad i = \overline{1, n}, \quad (3)$$

$$x_i \geq 0 \quad i = \overline{1, n}, \quad (4)$$

x_i - цілі .

Справді, з (3) і (4) випливає, що $x_i \leq 1$, а з (5) отримуємо $x_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad i = \overline{1, n}$. З останнього та

$\forall x_i = 1, x_j = 1: \quad a_{ij} = 0, \quad i, j = \overline{1, n}$ виходить (3), оскільки якщо $x_i = 0$, то $\sum_{j \neq i} a_{ij} x_j$ містить n-1 доданків і кожен доданок дорівнює 0 або 1, тобто $\sum_{j \neq i} a_{ij} x_j \leq n-1$; з іншого боку якщо $x_i = 1$, то

з умов $\forall x_i = 1, x_j = 1: \quad a_{ij} = 0, \quad i, j = \overline{1, n}$ випливає $\sum_{j \neq i} a_{ij} x_j = 0$ і (3) перетворюється у рівність.

Для розв'язування задачі (2)-(5) можна застосовувати відомі методи відсікання: Гоморі, Дальтона-Левіна, Фінкільштейна тощо. Отриманий вектор x визначає, які символи входять у слово, а аналіз положення їх дуг вказує порядок символів у слові. Тим самим алгоритм розпізнавання зображення разом написаного слова доведено до завершення.

Запропонований метод експериментально апробовано на ПК IBM PC 486 (80 М). Швидкість розпізнавання українськомовних текстів складає приблизно один символ за дві секунди. Час розпізнавання розподілюється майже рівномірно між трьома етапами.

Розпізнавання рукописних текстів шляхом аналізу топологічних елементів. Розпізнавання разом написаних слів шляхом послідовного та граматичного аналізу топологічних примітивів складається також із декількох етапів.

На етапі попереднього оброблення здійснюється виділення слів та вертикальна і горизонтальна нормалізації кутів нахилу. Виділення слів здійснюється так само, як і в попередній частині статті. Нормалізація кутів проводиться поворотами зображення текстів. Операції поворотів дискретного зображення призводять до появи додаткових збурень на контурах зображень текстів, тому для згладжування використовується фільтрація методом матричної маски.

На другому етапі здійснюється сегментація слів на конструктивні топологічні елементи. Такими елементами можуть бути окремі прості символи (*e, c, o...*), а також різні крючки, палочки тощо. При сегментації слово сканується у вертикальному і в горизонтальному напрямках. При скануванні у вертикальному напрямку будуються відповідні гістограми і обчислюються декілька числових параметрів: загальна висота зображення слова, положення центральної лінії зображення, наявність і положення центральної (символи *a, u, n...*), верхньої (для символів *v, b, A, B...*), нижньої смуг (символи *y, d, p...*). Ці параметри дозволяють розбити вхідний алфавіт на декілька підалфавітів ще на етапі сегментації. При горизонтальному скануванні будуються гістограми: розміщення зображення відносно середньої лінії; загальної висоти символів; кількості контурних пікселів; кількості контурів на лінії сканування.

Аналіз цих гістограм дає можливість розбити слово на окремі прості символи або конструктивні елементи.

Третій етап включає в себе розпізнавання окремих символів та елементів. При цьому зостосовуються два алгоритми - метод просторових областей та граматичний аналіз конфігурації контурів. Для послідовного аналізу при горизонтальному і вертикальному скануванні окремих символів будуються 3-, 5-, 7- вимірні вектори кількості пікселів в 3, 5, 7 рівномірних смугах зображення відповідно (рис. 5). Ці вектори нормалізуються, що зменшує залежність розпізнавання від розмірів символів. Таким чином, елемент відображається точками у відповідних 3-, 5-, 7- вимірних просторах. Для кожного можливого елемента у цих просторах задаються відповідні області (рис. 6).

Останні формуються у процесі настроювання або навчання системи як значення між максимальними та мінімальними параметрами компонент навчальної послідовності. Okрім цього, для кожного класу в процесі навчання будеться набір еталонних векторів. Це вектори середніх значень відовідних компонент навчальної послідовності. При розпізнаванні перевіряється, чи належить вектор елемента, що розпізнається, до області певного класу. Якщо він належить тільки одному класу (у всьому наборі просторів характеристик), то символ вважається розпізнаним. Якщо вектор належить декільком областям, то будеться відповідна послідовність гіпотез, які будуть перевірятися іншим методом. І коли вектор не належить жодному класу, то об'єкт вважається нерозпізнаним і для ідентифікації потрібний додатковий аналіз. У випадках

неоднозначного розпізнавання або нерозпізнання для додаткового аналізу використовуються алгоритми граматичного аналізу контурів. При цьому на зображені виділяються верхній та нижній півконтури. Кожний з них кодується за допомогою системи напрямлених відрізків Фрімана, представляється у вигляді орієнтованого графа та розпізнається методом порівняння з еталоном [3,7]. Для кожного класу отримують два еталонних графи. Еталонний граф представлений у вигляді регулярної програмної граматики, а розпізнавання вхідного символу полягає у граматичному аналізі відповідного вхідного графа. Для впорядкування гіпотез зastosовується метод найближчого сусіда, тобто враховується відстань між вектором характеризуючих компонент поточного символа і еталонними векторами середніх значень. Впорядкування гіпотез суттєво впливає на швидкість розпізнавання конструктивних елементів.

На останньому етапі з конструктивних елементів складаються символи і слова, а ідентифікація слів зводиться до оптимізаційної задачі з максимізацією кількості залучених конструктивних елементів.

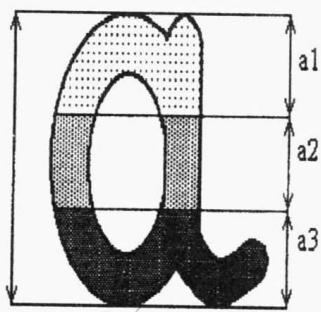


Рис. 5.

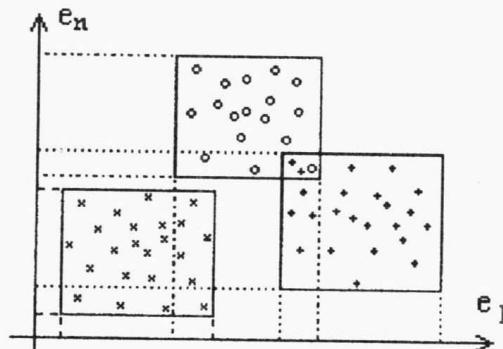


Рис. 6.

Якість і швидкість розпізнавання на експериментальних текстах адекватні до параметрів первого методу.

Література

- Гайдайчук І., Фратавчан В. Граматичні методи розпізнавання рукописних текстів під час сканерного введення //Україномовне програмне забезпечення. Матеріали 4-ої та 5-ої Міжнародних науково-практичних конференцій "УКРСОФТ". - Львів, СП "БаK", 1995, с. 195-199.
- Фор А. Восприятие и распознавание образов. - Москва: Машиностроение, 1989, 271 с.
- Фу К. Структурные методы в распознавании образов. - Москва: Мир, 1977, 320 с.
- Фу К. Последовательные методы в распознавании образов. - Москва: Мир, 1976, 255 с.
- Фу К., Гонсалес Р., Ли К. Робототехника. - Москва: Мир, 1989, 624 с.
- Хорн Б.К.П. Зрение роботов. - Москва: Мир, 1989, 487 с.
- Fratavchan V., Gaidaichuk I., Rusnak M. Matrix and Grammatical Methods for Pattern Recognition // Development and Application Systems, No. 4, Suceava, 1994, pp. 203-208.
- Fratavchan Valery, Gaidaichuk Igor, Rusnak Mykola Application of Structure Analysis for Manuscripts Recognition // Development and Application Systems, Suceava, Romania, 1996, Number 7, pp. 125-130.

Технологія вимірювання просторової роздільної здатності інфрачервоного давача, яка базується на автоматичному розпізнаванні тест-об'єкту

МИХАЙЛО ПОПОВ, СЕРГІЙ БАЛАШОВ

Інститут військово-повітряних сил

252186 Київ, просп. Повітрофлотський 30
Тел.: (044) 271-6418

Михаил Попов, Сергей Балашов. Технология измерения пространственной разрешающей способности инфракрасного датчика, которая базируется на распознавания тест-объекта.

Описывается технология измерения пространственной разрешающей способности инфракрасного датчика.

Просторова роздільна здатність (ПРЗ) є найважливішою характеристикою інфрачервоного (ІЧ) давача, знання якої дозволяє не тільки оцінити якість зображення, що формується, але і спрогнозувати вірогідність виявлення і можливість розпізнавання об'єктів, які знаходяться в полі

зору давача. При вимірюванні ПРЗ звичайно використовуються тест-об'єкти у вигляді мір, які складаються із почергових теплих і холодних смуг, температура окремої із них підтримується незмінною, а ширина смуги поступово змінюється за встановленим законом [1]. Інфрачервоне зображення такої міри являє собою послідовність чорно-білих пар смуг співвідносної ширини, кожній із котрих надається свій порядковий номер n : $n=1,2,\dots,N$; N - загальне число пар смуг міри.

При аналізі зображення через відомі чинники видимість смуг, починаючи з деякої, обмежена. Номер останньої із пар смуг, що ще сприймаються на зображені, і є ПРЗ; будемо позначати цей номер через n_r .

У відомих методиках вимірювання ПРЗ спостереження і розрізнення штрихів міри на зображені виконуються людиною-десифровщиком, яка і видає розмір n_r . Наукові дослідження показують, що оцінка цієї величини достатньо різиться навіть в межах колективу десифровщиків однієї кваліфікації. Детальний аналіз впливу суб'єктивного фактору на процес розрізнення штрихів тест-об'єкту викладені в [2].

Проведені дослідження дають підставу разглядати n_r як сходинкову величину, яка підпорядковується нормальному закону розподілу. Нехай μ - квантиль розподілу, s - середньоквадратичне відхилення, e - заданий певний інтервал. Тоді відомо [3], що для отримання спроможної і достовірної оцінки n_r потрібно $(us/e)^2$ десифрувальників, тобто приблизно декілька десятків людей - спеціалістів достатньо високої кваліфікації. Виконати цю умову на практиці досить проблематично, тому нами розроблена технологія вимірювання ПРЗ ІЧ датчика, яка базується на автоматичному розпізнаванні тест-об'єктів.

Сутність технології можна відобразити за допомогою такого алгоритму: попереднє оброблення отриманого зображення з метою відокремлення завад; розрахунок параметрів Хо (Hough) і порогове оброблення отриманої площини параметрів; пошук інформативних ознак у площині параметрів (локальних максимумів, розташованих у певному порядку) і приняття рішення про наявність тест-об'єкту на зображені; перехід у площину зображення, з одноразовим обчисленням координат тест-об'єкту, прийняття рішення про придатність зображення до подальшого оброблення; за умови виконання попереднього пункту проводиться аналіз зображення тест-об'єкту і знаходитьться група смуг, яка знаходиться на межі сприйняття; знаходження ПРЗ з урахуванням априорного знання параметрів тест-об'єкту.

Запропонована технологія вимірювання ПРЗ була перевірена експериментальними дослідженнями і дала позитивні результати. Для перевірки були використані матеріали тренувальної сертифікації, яка були отримані в Льотному Центрі ВПС Росії в смт. Шаталово (1994 рік).

Реалізація цього алгоритму можлива на сучасній обчислювальній техніці.

Так, використання персональних комп'ютерів IBM AT з процесором 486DX4-100 дозволяє робити необхідні обчислення майже в реальному часі. Практичне використання цієї технології можливе в рамках Договору з відкритого неба, в якому Україна активно співпрацює з іншими державами, починаючи з 1992 року. Крім того, використання цієї технології в період випробувань ІЧ техніки, призначеної для екологічного моніторингу, картографічної зйомки місцевості, а також у військових цілях, значно підвищить якість систем, що проектуються.

Література

1. Ллойд Дж. Системы тепловидения: Пер. с англ. - Москва, Мир, 1978, 414 с.
2. Балашов С.В. О влиянии субъективного фактора на процесс дешифрирования тест-объектов // В сб. научных трудов "Информационные технологии в дешифрировании изображений. Классификация и оценка эффективности". - Киев: МО Украины, 1996, с. 8-13.
3. Алферов В.В., Мезенцев В.Г. Обработка опытных данных и представление результатов измерений - Москва, Воениздат, 1979, 237 с.



The Linear Models of Optimal Synthesis of the Discrete Zonal Images

SERHIJ STANKEVYCH

Air Force Institute

252186 Kyiv, Povitroflets'kyj prospekt 30
Phone: (044) 271-6236

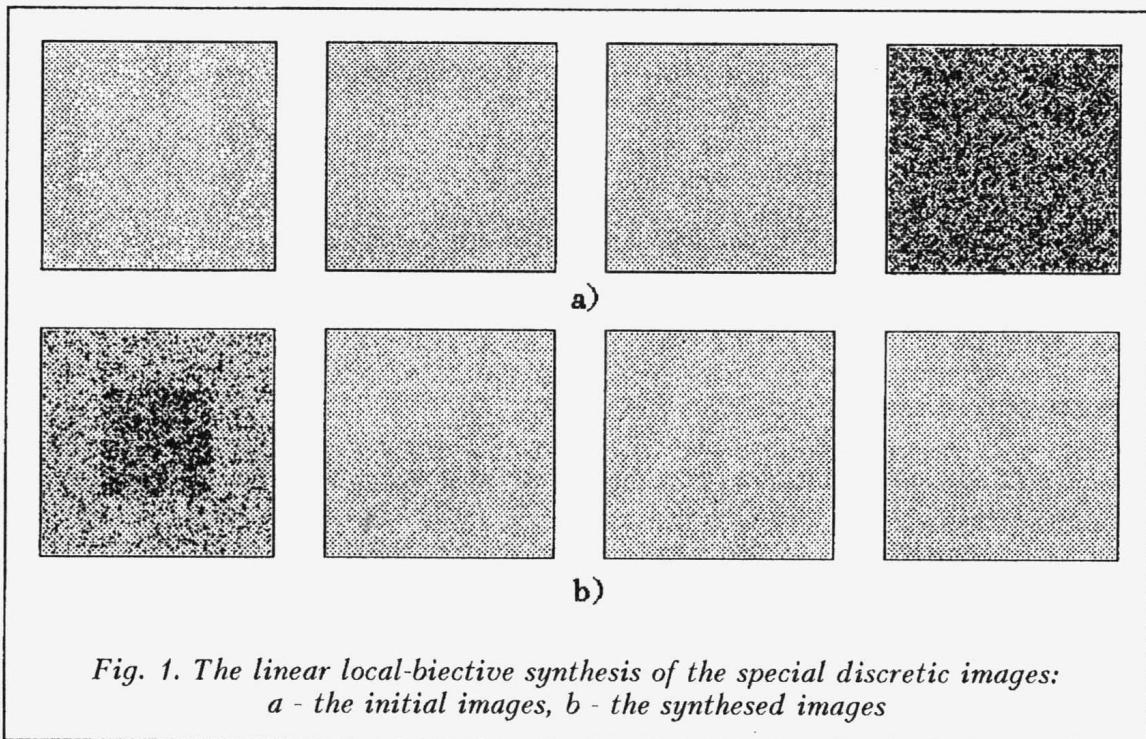
Сергій Станкевич. Лінійні моделі оптимального синтезу дискретних зональних зображень.

Розглянута адаптація відомих лінійних моделей статистичного аналізу до розв'язання задачі синтезу дискретних зображень за матеріалами багатозонального аерокосмічного моніторингу.

The processes of synthesis of the digital images has a greatest importance for the automated data processing of aerospace monitoring of an Earth surface. In many cases the synthesis of the complete set of spectrozonal images is a unique way to grant identification features of objects of aerospace monitoring to human-operator for their complex interpretation.

Hereinafter the synthesis of the discrete images is understood as a forming of some new discrete image by joint processing of several initial images. Let Υ be the set of the discrete images, each of which $I \in \Upsilon$ is described by spatial distribution of intensity x_i , $i \in I$ of raster's elements (pixels). Then the synthesised image J can be submitted as $J = \mathfrak{R}\{\Upsilon\}$, where $\mathfrak{R}\{\cdot\}$ is some generalized operator of synthesis. In practice usually action of the operator $\mathfrak{R}\{\cdot\}$ is limited in the field of spatial-corresponding pixels of the initial images (local operator of synthesis) or of the members of Set Theory relation $\mathfrak{R}\left\{\prod_{I \in \Upsilon} I\right\}$ (the local-bijective operator, if $\mathfrak{R}\{\cdot\}$ is a bijective).

It is necessary to consider as a main task of synthesis the concentration of information, that contained in the complete set of the initial zonal images with the purpose of facilitation of



*Fig. 1. The linear local-bijective synthesis of the special discrete images:
a - the initial images, b - the synthesised images*

opportunities of its perception and recognition. The basic opportunity of the decision of the given problem leans on hypothesis of existence of the directly non-observed synthesised image, that reflect some immanent properties of registered objects and ensure minimum losses of the information.

The linear local-bijective model of the synthesis of discrete images assumes a construction of the complete set \mathfrak{J} of the synthesised images $J \in \mathfrak{J}$, dimension of which, as a rule, is much less than at complete set of the initial images, looks as

$$Y = F X + C, \quad (1)$$

where $X = \{X(I)\}$, $I \in \Upsilon$ is a vector of intensity of the complete set of the initial images, $Y = \{Y(J)\}$, $J \in \mathfrak{J}$ is a vector of intensity of the complete set of the synthesised images, F is a matrix of linear transformation with $|\Upsilon| \times |\mathfrak{J}|$ dimension, C is a $|\mathfrak{J}|$ -dimensional vector of the intensity corrections. It is a traditional local linear model, in which intensities of the corresponding zonal images are represented as $X(I) = S(I) + U(I)$, $\forall I \in \Upsilon$, where $S(I)$ is a useful signal into the zonal image I , $U(I)$ is a corresponding zonal handicap.

Imposing of additional norming limitations to elements of a matrix F and choice of the corrections C by condition of the coordination of dynamic bands of the synthesised images and their interpreter results to common for all considered models a way of their determination: each line of a matrix F is a corresponding eigenvector of some auxiliary matrix Ψ , the kind of which is defined by used model. The Fig. 1 displays a results of the synthesis of four special digital images according to (1).

According to dependence by chosen criterion of optimum it can be allocated the following linear models of synthesis:

1. The model of a principal components [1]. It provides a maximization of an intensity total variance of complete set of the synthesised images. A matrix Ψ is a covariance matrix of the initial images: $\Psi = \text{Cov}_{I \in Y} X(I)$. Largest informativity has the synthesised image that corresponding to the largest eigenvalue of a matrix Ψ .

2. The model that maximize a "signal / noise" ratio in the synthesised image [2]. The matrix Ψ looks as $\Psi = SV^{-1}S^T$, where $S = \{S(I)\}$, $I \in Y$ is a vector of zonal signals, $V = \text{Cov}_{I \in Y} U(I)$ - a covariance matrix of noise of the zonal images.

3. The model, that minimize the average weighed losses of the Fisher's information. The matrix Ψ is defined [3] as $\Psi = V^{-1/2} \left[\int_{i \in I} \partial S / \partial i (\partial S / \partial i)^T p(i) di \right] V^{-1/2}$, where $\partial S / \partial i$ is a vector of derivative of zonal signals on area of the corresponding images, $p(i)$ is a distribution of probability of a signal site. In particular, when $p(i) = \text{Const}$: $\Psi = V^{-1/2} \left[1/|I| \sum_{i \in I} \partial S / \partial i (\partial S / \partial i)^T \right] V^{-1/2}$.

4. The model of the Factor Analysis. The vector of optical signals of the initial images is represented as $X = WY + U$, where Y is a set of the latent factor images, U is a residual stochastic component (vector of deviations), W is a matrix of a factor's weights of $|X| \times |Y|$ dimension. The synthesised image can be treated as a reflection of the objective factor that induced the complete set of the observable zonal images $X(I)$, $I \in Y$. From Main Factor Theorem [4] follows, that $\text{Cov}_{I \in Y} X(I) = WW^T + V$.

The matrix W is a set of first $|Y|$ eigenvectors of a matrix Ψ . The matrixes Ψ and V can be determined by means of consecutive iterations: $\Psi^{(0)} = \text{Cov}_{I \in Y} X(I) - V^{(0)}$, $V^{(0)} = \left[\text{Diag} \left[\text{Cov}_{I \in Y} X(I) \right]^{-1} \right]$ is a zero approximation of matrixes Ψ and V , $\Psi^{(1)} = W^{(0)} [W^{(0)}]^T$, $V^{(1)} = \text{Cov}_{I \in Y} X(I) - \Psi^{(1)}$ is a first approximation of matrixes Ψ and V etc. The iterative process is finished when it is accomplished a condition $\text{Tr Diag} [\text{Cov}_{I \in Y} X(I) - W^{(n)} [W^{(n)}]^T - V^{(n)}] < 0.005$.

Further the complete set of the synthesised images can be determined by one of the known methods of the Factor Analysis, for example, by M. Barlett's method: $Y = [W^T V^{-1} W]^{-1} W^T V^{-1} X$.

The considered linear models of synthesis can be used in two different variants: a statistical models, in which the calculation of the initial characteristics of the complete set of the images is executed on the whole area, and a local models, in which the calculation is executed within the frames of some scanning window. In the latter case the synthesis procedure becomes more flexible due to obtaining of the additional information about the image segments, size of which is more than window size. Thus it is obviously essential increasing of required quantity of calculations. The dependence of quality of the local synthesised image from relative size of a scanning window is illustrated by Fig. 2.

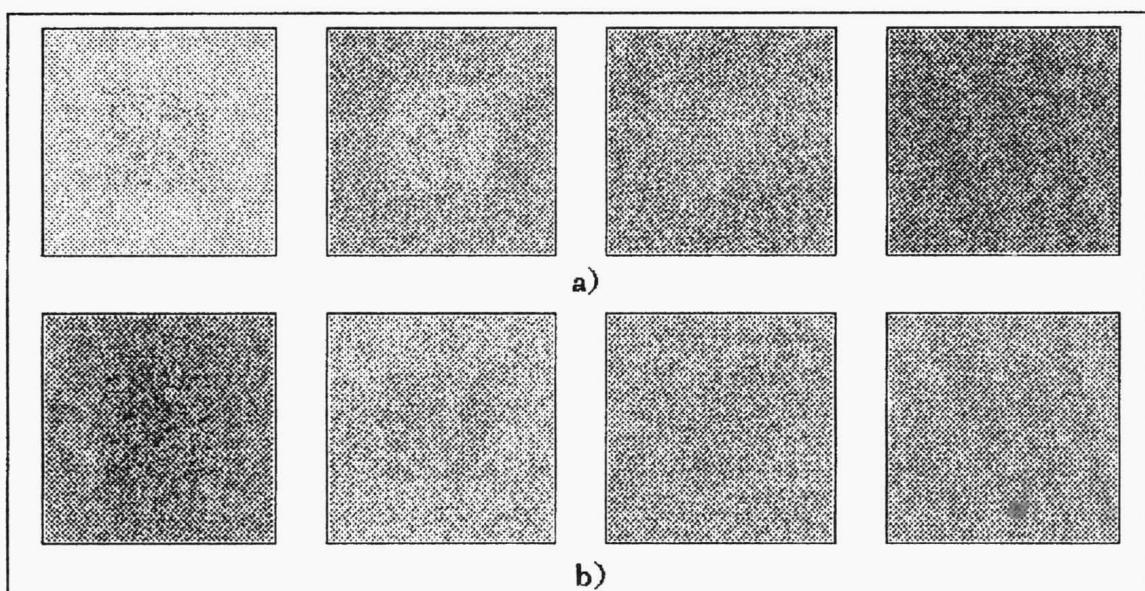


Fig. 2. Dependence of actual contrast of the synthesised image from relative size of a scanning window

Thus, given models permit rather simply to design and to modify the local-bijective operator of optimal synthesis of the complete set of discrete images, that is important at processing of landscape aerospace monitoring data, that accomplished simultaneously at several zones of electromagnetic spectrum.

References

1. Arbuzov P.A., Stankevich S.A. About Method of Spectral Reflection Coefficient Using in Object Recognition in Remote Sensing // Proceedings of the First All-Ukrainian Conference "Signal/Image Processing and Pattern Recognition". - Kyiv, 1992, p. 154.
2. Левин Б.Р. Теоретические основы статистической радиотехники.- Москва: Радио и связь, 1989, 656 с.
3. Latyshev V.V., Vinokurov S.N. Feature Extraction in the Presence of Parameter Uncertainty // Proceedings of the Second All-Ukrainian Conference "Signal/Image Processing and Pattern Recognition". - Kyiv, 1992, p. 29-33.
4. Харман Г. Современный факторный анализ / Пер. с англ. - Москва: Статистика, 1972, 486 с.



The Grammatical Approach to Documents Analysis: Applied to Summaries

S. TAYEB-BEY, S. SAIDI, H. EMPTOZ

INSA de Lyon

Reconnaissance de Formes & Vision Laboratory

20 Avenue Albert Einstein, 69621 Villeurbanne Cedex

Tel : (33) 72 43 80 93, fax : (33) 72 43 80 97

E-mail : tayebbe@rfv.insa-lyon.fr

S. Tayeb-Bey, S. Saidi, H. Emptoz. The Grammatical Approach to Documents Analysis: Applied to Summaries

This paper deals with the use of grammatical formalism to recognize the physical and logical structures of a composite document.

A document has three principal characteristics:

- ① its content,
- ② its logical structure, i.e. its layout and the logical organization of its elements of information,
- ③ its physical structure, i.e. the position of information elements on the pages.

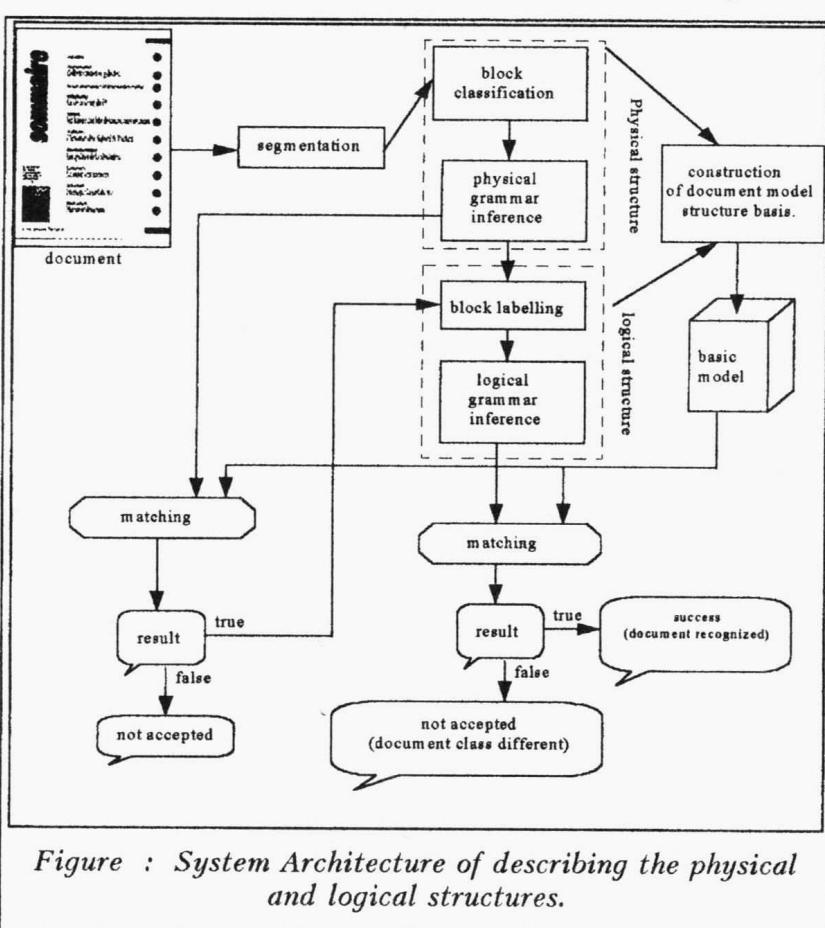
A lot of work has been done in document structuring, showing the growing interest in this field in recent years. As result, several methods have been proposed to solve the problem of structure representation.

Ingold [1] proposed a document description language similar to an attributed grammar. The aim in designing this language has been to top-down analyse several document classes.

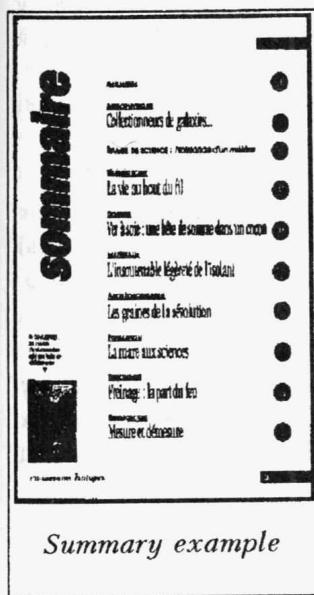
Chenevoy [2,3] proposed a general-purpose system for document structure recognition called GRAPHIEN. The system organizes and controls the diverse document recognition processes.

However, the construction of structure models of documents has turned out to be a difficult task, and is often carried out manually because of document diversities.

Grammatical formalism is one of the approaches used in syntactic pattern recognition and it is quite suitable for document analysis. This formalism has also turned out to be a powerful tool in describing a document.



We are interested in structuring documents like summaries, the aim of our research is to create a document structuring system (see figure) by using a two level grammar [4]. Work on this system is currently under way.



Two level Grammars. A two level grammar, also W-grammar [5] is a formal system well adapted to the language definition of. It is composed of two grammars called metagrammar and hypergrammar. The metagrammar ($\text{type} \geq 2$) defines the possible domain of value for metavariables. These metavariables appear in the rules of the hypergrammar.

By applying the principle of uniform replacement in the hyperrules (similar metavariables are replaced by the same value), we obtain a ground instance of the hypergrammar called the protogrammar containing only context-free rules. Note that the protogrammar may potentially be infinite.

The system. We propose a new system for document recognition and analysis. Prior to any analysis, we start with:

A learning system concept: the purpose of this step is to infer a grammar of the physical structure called physical grammar , and then to infer a grammar of the logical structure called logical grammar. The result of this step is to construct a model of basic document structure.

Then, in the **document recognition** step: the system compares a specific document summary with the model of the learning system. If this document is not recognized as a known summary, the physical and logical grammar will be updated by adding new rules.

The system uses W-grammars in which the physical and logical grammars are given in the metagrammar. Then, the hypergrammar will describe the transformation of the physical and logical structures among other calculus.

In our system, we use an operational version of W-grammar (called transparent W-grammar).

References

1. R. Ingold. A Document Description Language to Drive Document Analysis. ICDAR-91, Vol. 1, pp. 294-301.
2. A. Belaid, J. J. Brault and Y. Chenevroy. Knowledge-Based System for Structured Document Recognition. In MVA'90 IAPR Workshop on Machine Vision Applications, November 1990.
3. Y. Chenevroy. Reconnaissance structurelle de documents imprimés : Etudes et Réalisations. Ph.D. Thesis, INRIA-Lorraine, December 1992.
4. S. Saidi. Extensions Grammaticales de la Programmation (en) Logique : Application à la Validation des Grammaires Affixes, Ph.D. Thesis, Ecole Centrale de Lyon, 1992.
5. A. A. Van Wijngaarden. *Orthogonal Design and Description of Formal Languages*. Mathematisch Centrum Amsterdam, MR 76, 1965.



Ефективна сегментація зображення в процесі його дворазового перегляду

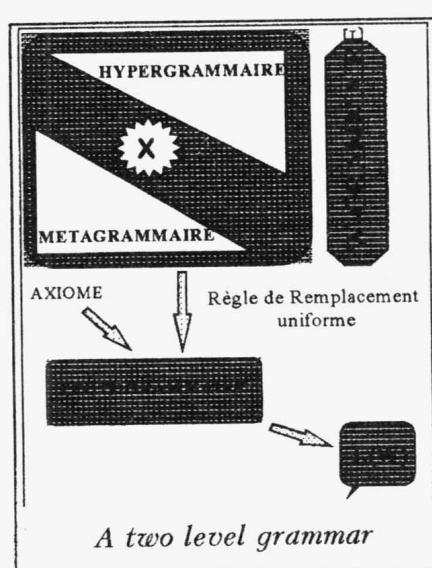
ДМИТРО ШЛЕЗІНГЕР

Національний технічний університет - КПІ

252056 Київ, просп. Перемоги 37

Дмитрий Шлезингер. Эффективная сегментация изображения в процессе его двукратного просмотра.

Описан алгоритм разметки связных компонент изображения. Учет специфических свойств изображения как объекта обработки, позволяет построить алгоритм, вычислительно более эффективный, чем алгоритмы решения более общей задачи разметки связных компонент неориентированных графов. Описанный алгоритм решает задачу за время, линейно зависящее от размера изображения, не требуя при этом никакой дополнительной памяти.



A two level grammar

Вступ. Відомо, що задача виділення сполучних компонентів неорієнтованого графа вирішується алгоритмом, кількість операцій якого пропорційна $(m+n)$, де n - кількість вершин графу, а m - кількість ребер [1]. Цей алгоритм потребує допоміжну пам'ять, обсяг якої пропорційний n . Лишається відкритим питання про існування алгоритму, що вирішує цю задачу за час, що лінійно залежить від обсягу графа, і не потребує ніякої допоміжної пам'яті. У цій праці описано алгоритм, що вирішує задачу лише для графів, що описують зображення. Особливості зображення як об'єкту оброблення дозволяють конструювати алгоритми, оптимальні як за швидкодією, так і за пам'яттю.

Формулювання задачі та аналіз відомих рішень. Нехай V - скінчenna множина, $R \subset V \times V$ - симетричне відношення сусідства, а R^* - рефлексивно-транзитивне замикання R (відношення еквівалентності). Пара (V, R) є неорієнтованим графом, а класи еквівалентності R^* - це сполучні компоненти цього графа. Сполучні компоненти звичайно описуються цілочисельною функцією ϕ , заданою на множині вершин, такою, що $(\phi(v_1)=\phi(v_2)) \Leftrightarrow ((v_1, v_2) \in R^*)$ для всіх $v_1, v_2 \in V$. Задача полягає в побудові алгоритму, який для будь-якого графа (V, R) буде функцією ϕ .

Алгоритми, що не потребують допоміжної пам'яті, вирішують задачу в два етапи. На першому етапі обчислюються ті або інші дані, які розташовуються в пам'яті, призначений для остаточних результатів. Ці проміжні дані є такими, що вони можуть бути перетворені в остаточні в процесі одноразового перегляду масиву ϕ . Це перетворення є змістом другого етапу алгоритма.

Розглянемо докладніше ці ідеї.

Нехай множина V вершин початкового неорієнтованого графу - множина цілих чисел. D -подання графа, (V, R) - це орієнтований граф із такими властивостями: з будь-якої вершини виходить одне ребро; сполучні компоненти цього орієнтованого графа тотожні сполучним компонентам графа (V, R) ; в кожній сполучній компоненті є вершина, що називається коренем; ребро, що виходить з неї, до неї ж і входить.

D -подання графа (V, R) описується функцією $\phi: V \rightarrow V$, де $\phi(v)$ - вершина, до якої входить ребро, що виходить з v . Побудування сполучних компонентів графа (V, R) за його D -поданням здійснюється такою простою програмою:

```
ДЛЯ i=1, 2, ..., n
{ j=i ; ДОКИ (φ(j)≠j)   j=φ(j) ;
  k=i ; ДОКИ (φ(k) ≠ j) { k'=k ; k=φ(k) ; φ(k')=j ; }
}
.
```

Для конструювання алгоритму побудови D -подання графа (V, R) розглянемо функцію $\phi^*: V \rightarrow V$ для будь-якої заданої функції $\phi: V \rightarrow V$, що визначається такими правилами: якщо $\phi(v)=v$, то $\phi^*(v)=v$, інакше $\phi^*(v)=\phi(\phi(v))$. В подальшому оператор $u=\phi^*(v)$ означає таке:

$u=v$; ДОКИ $(\phi(u) \neq u)$ $u=\phi(u)$.

Алгоритм побудови D -подання графу (V, R) має вигляд:

ДЛЯ $i=1, 2, \dots, n$ $\phi(i)=i$

ДЛЯ $i=1, 2, \dots, n$

ДЛЯ КОЖНОГО j , ТАКОГО, ЩО $(i, j) \in R$ ТА $j < i$, $\phi(\phi^*(j))=i$.

Складність цього алгоритму має порядок $(n+m) \times n$.

Наведений алгоритм буде такі D -подання графа (V, R) , що для будь-якої вершини i справджується $\phi(i) \geq i$. Це дає можливість побудувати алгоритм перетворення D -подання в сполучні компоненти значно простіше, ніж це було представлено раніше, а саме так:

ДЛЯ $i = n, n-1, \dots, 1$ $\phi(i) = \phi(\phi(i))$.

В тих випадках, коли граф (V, R) описує зображення, сполучні компоненти можуть бути виявлені набагато швидше.

Виявлення сполучних компонентів зображення. Нехай T - прямокутна ділянка двовимірної цілочисельної гратки розміром $m \times n$. $T = \{(i, j) | 1 \leq i \leq m, 1 \leq j \leq n\}$. Зображення - це функція $v: T \rightarrow \{0, 1\}$, якій відповідає граф (V, R) , де $V = \{(i, j) | v(i, j) = 1\}$, а $R = \{((i, j), (i', j')) | v(i, j) = v(i', j') = 1, |i-i'| + |j-j'| = 1\}$. Задача полягає в побудові алгоритму, що для будь-якого зображення v буде функцію $\phi: T \rightarrow T$, таку що $(\phi(i, j) = \phi(i', j')) \Leftrightarrow ((i, j), (i', j')) \in R^*$, де R^* - рефлексивно-транзитивне замикання відношення R .

Алгоритм має складатися з двох частин, як і алгоритм для довільних графів: побудова D -подання зображення та перетворення D -подання в сполучні компоненти. Як D -подання, так і сполучні компоненти будемо описувати функцією $\phi: T \rightarrow T$, при цьому $\phi(i, j) = 0$ означає, що точка (i, j) - біла, тобто не належить жодній сполучній компоненті.

Друга частина алгоритму така ж, як і для довільних графів, а саме:

ДЛЯ $i = m, m-1, \dots, 1$

ДЛЯ $j = n, n-1, \dots, 1 \quad \phi(i,j) = \phi(\phi(i,j))$

Усі труднощі рішення зосереджені в першій частині, тобто в побудові **D**-подання. Для її конструювання визначимо необхідні допоміжні поняття.

Визначимо на множині пікселів зображення відношення порядку, який будемо називати порядком перегляду зображення; з двох пікселів (i,j) і (i',j') , розташованих в різних рядках ($i < i'$), раніше переглядається піксель (i,j) з меншим номером рядка, а з двох пікселів, (i,j) і (i,j') , ($j < j'$), розташованих в одному рядку, раніше переглядається піксель (i,j) з меншим номером j . Нехай i^*, j^* - координати останнього переглянутого піксела. Сукупність пікселів (i,j) , де $i < i^*$ або $(i=i^*) \wedge (j < j^*)$, будемо називати переглянутою частиною зображення, сукупність пікселів (i,j) , де $(i=i^*) \wedge (j < j^*)$ або $(i=i^*-1) \wedge (j=j^*)$, - її межею, а сукупність усіх інших переглянутих пікселів – її внутрішньою частиною.

Стан алгоритму побудови **D**-подання в будь-який момент його виконання характеризується функцією $\phi: T \rightarrow T$, яка на різних стадіях має різну змістовну суть. Після перегляду усіх рядків зображення ця функція має містити **D**-подання такого вигляду.

Вимога 1. Для будь-якого чорного піксела з координатами (i,j) $\phi(i,j)$ визначає координату i^*, j^* , що має такі властивості:

а) піксель i^*, j^* переглядається при прийнятому порядку перегляду після піксела i, j ;

б) піксель i^*, j^* належить тій самій сполучній компоненті, що піксель i, j ;

в) піксель i^*, j^* в прийнятому порядку перегляду розташований раніше будь-якого іншого піксела, що має властивості а) та б);

г) якщо піксела з властивостями а), б) та в) не існує, то $\phi(i,j)=i, j$.

Для будь-якого білого піксела з координатами (i,j) $\phi(i,j)=(0,0)$.

Нехай на якійсь стадії переглянуто повністю i^* рядків. На цьому етапі функція ϕ має задовольняти такі вимоги.

Вимога 2. а) Для усіх пікселів усередині переглянутої частини функція ϕ має задовольняти вимогу 1.

б) Нехай (i^*, j) - чорний піксел на межі переглянутої частини, $(i^*, j+1)$ - сусідній з ним білий піксель, а (i^*, j') найближчий до піксела (i^*, j) чорний піксель, зв'язаний з (i^*, j) на переглянутій частині, при цьому $j' > j$. У такому разі $\phi(i^*, j)=(i^*, j')$.

в) Якщо (i^*, j) - чорний піксел на межі переглянутої частини, а піксель (i^*, j') , що задовольняє вимоги п. б), не існує, то $\phi(i^*, j)$ дорівнює (i^*, j'') , де j'' - мінімальна координата чорного піксела, зв'язаного на переглянутій частині з пікселом (i^*, j) .

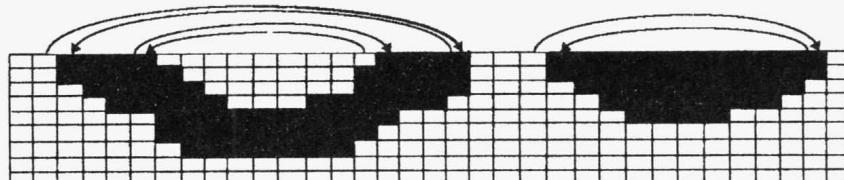
г) Нехай (i^*, j) - білий піксель на межі переглянутої частини, до того ж такий, що піксел $(i^*, j+1)$ - чорний. Нехай (i^*, j') - такий піксел, що $\phi(i^*, j')=(i^*, j+1)$. В такому разі для білого піксела (i^*, j) функція ϕ має дорівнювати (i^*, j') .

д) Для усіх інших пікселів на межі переглянутої частини значення ϕ дорівнює $(0,0)$.

Вимога 2 ілюструється рисунком, де у вигляді стрілок зображена функція ϕ для межі переглянутої частини. З рисунка зрозуміла змістовна суть значень ϕ на межі переглянутої частини. А саме, стрілка, що виходить з чорного піксела, скажімо (i^*, j) , праворуч від якого знаходиться білий, вказує на координату чорного піксела, що знаходиться на межі і примикає до другого кінця контуру, зв'язаного на переглянутій частині зображення з (i^*, j) . Такий самий зміст має стрілка, що виходить з білого піксела на межі, справа від якої знаходиться чорний піксель.

Зрозуміло, що якщо вимога 2 виконується після перегляду кожного рядка, то після перегляду останнього рядка виконується і вимога 1, якщо тільки останній рядок складається лише з білих пікселів.

Програма оброблення рядка має бути створена так, щоб після перегляду кожного піксела зберігалася змістовна суть стрілок, що виходять з пікселів на межі, що примикають до контура. До того ж, якщо контур проходить між пікселями (i^*, j^*) та (i^*-1, j^*) , а другий кінець його проходить між пікселями (i, j) та $(i, j+1)$, функція $\phi(i, j)$ має дорівнювати деякому особливому значенню, що не дорівнює ніякій парі координат (i, j) (визначимо це значення *). Координата другого кінця контуру, перший кінець якого проходить між пікселями (i^*, j^*) та (i^*-1, j^*) , зберігається в допоміжній парі комірок, яку визначимо як i_0, j_0 .



Алгоритм оброблення i^* -го рядка працює кроками, номери яких визначені як $j=2, \dots, m$. На j -ому кроці обробляється четвірка пікселів з координатами $(i^*, j), (i^*, j-1), (i^*-1, j)$ та $(i^*-1, j-1)$. В залежності від чорнот цих пікселів виконуються ті або інші дії, що змінюють пару i_0, j_0 та значення ϕ в тих або інших пікселях так, що на кожнім кроці алгоритму стрілки на межі мають вказаний зміст. Далі в таблиці вказані дії, які мають бути виконані при кожній з 16-ти можливих комбінацій чорнот.

- $(i_0, j_0) = (i, j);$
 $\phi(i, j-1) = *.$
- Жодних дій.**
- Якщо $\phi(i-1, j-1) \neq *$
{ $\phi(\phi(i-1, j-1)-1) = (i_0, j_0);$
 $\phi(i_0, j_0) = \phi(i-1, j-1);$ }
Інакше $\phi(i-1, j-1) = (i-1, j-1).$
- Виконати  , потім  .
- Виконати  , потім  .
- Жодних дій.**
- $\phi(i, j-1) = (i, j).$
- Якщо $\phi(i-1, j-1) \neq *$
{ $\phi(i_0, j_0-1) = \phi(i-1, j-1);$
 $\phi(\phi(i-1, j-1)) = (i_0, j_0);$ }
- $\phi(i-1, j-1) = (0, 0).$
 $\phi(i_0, j_0) = (i, j);$
 $\phi(i, j-1) = (i_0, j_0);$
 $\phi(i-1, j-1) = (i-1, j).$

- $(i_0, j_0) = \phi(i-1, j-1);$
 $\phi(\phi(i-1, j-1)) = *;$
 $\phi(i-1, j-1) = (0, 0).$
- $(i_0, j_0) = (i, j-1);$
 $\phi(i, j-1) = *;$
 $\phi(i-1, j-1) = (i-1, j).$
- $\phi(i, j-1) = \phi(i-1, j-1);$
 $\phi(\phi(i-1, j-1)) = (i, j);$
 $\phi(i-1, j-1) = (0, 0).$
- $(i_0, j_0) = \phi(i-1, j-1);$
 $\phi(\phi(i-1, j-1)-1) = *.$
- $\phi(i, j-1) = \phi(i-1, j-1);$
 $\phi(\phi(i-1, j-1)-1) = (i, j-1).$
- $\phi(i, j-1) = (i_0, j_0);$
 $\phi(i_0, j_0-1) = (i, j-1).$
- $\phi(i-1, j-1) = (i-1, j).$

Значення $(\phi(i, j)-1)$ треба розуміти як $(i', j'-1)$, де $(i', j') = \phi(i, j)$.

В цілому алгоритм виявлення сполучних компонентів зображення має вигляд:

ДЛЯ УСІХ $i=1, \dots, m, j=1, \dots, n \quad \phi(i, j) = (0, 0);$

ДЛЯ УСІХ $i=2, \dots, m$

ДЛЯ УСІХ $j=2, \dots, n$

{ ОБРОБКА ЧЕТВІРКИ ПІКСЕЛІВ З КООРДИНАТАМИ $(i, j), (i, j-1), (i-1, j), (i-1, j-1)$ ЗГІДНО З ТАБЛИЦЕЮ; }

$K=1;$

ДЛЯ $i=m, m-1, \dots, 2$

ДЛЯ $j=n, n-1, \dots, 2$

{ ЯКЩО $(\phi(i, j) = (i, j))$, ТО $\{\phi(i, j) = K; K=K+1\}$

ІНАКШЕ $\phi(i, j) = \phi(\phi(i, j))$ }

Швидкодія даного алгоритму має порядок $n \times m$, де m - кількість рядків, а n - кількість пікселів у рядку зображення.

Література

- Ю.В.Капітонова, А.А.Летичевский. Математическая теория проектирования вычислительных систем.- Москва: Наука, 1988, 296 с.

