

基于边的图文法产生式选择无关条件的判断

王毅*, 丁函

(湖北文理学院 数学与计算机科学学院, 湖北 襄阳 441053)

(* 通信作者电子邮箱 dhnats@163.com)

摘要: 为了降低归约算法的时间复杂度, 在基于边的上下文相关图文法(EGG)形式化的基础上, 通过对产生式形式的适当约束, 提出了 EGG 的产生式选择无关条件的判断方法。通过此方法可有效判断 EGG 产生式的选择无关性。对于选择无关的产生式, 由于归约过程中产生式的使用顺序不会影响归约的结果, 从而避免了回溯, 能够有效地降低归约算法的时间复杂度。

关键词: 选择无关; 归约算法; 时间复杂度; 基于边的上下文相关图文法

中图分类号: TP391.4 **文献标志码:** A

Selection-free judgment for productions based on edge-based context-sensitive graph grammar

WANG Yi*, DING Han

(School of Mathematical and Computer Science, Hubei University of Arts and Science, Xiangyang Hubei 441053, China)

Abstract: In order to reduce the time complexity of parsing algorithm, based on the Edge-based context-sensitive Graph Grammar (EGG), a selection-free judgment method was proposed for the productions of EGG with appropriate constraints. This method can effectively judge the selection independence of the productions of EGG. For the selection-free productions, using the order of the productions will not affect the result of the parsing, thus the parsing process can avoid backtracking and the time complexity of the parsing algorithm can be effectively reduced.

Key words: selection-free; parsing algorithm; time-complexity; Edge-based context-sensitive Graph Grammar (EGG)

0 引言

图文法^[1]是一维字符文法在二维对象上的自然扩展, 为图语言的定义、图生成器及图分析器的设计提供理论基础。经过几十年以来不断的探索和研究^[2-5], 图文法理论已广泛应用于计算机学科各个领域^[6-8]。

一个图文法包含三个方面的元素: 文法的初始图、一组用于图的转换的产生式和嵌入规则。目前已存在很多图文法的形式化方法, 主要可以分为上下文无关^[9]和上下文相关两类。由于后者较前者有更强的表达能力, 所以得到了更广泛的应用。典型的上下文相关图文法有: CMG^[10] (Constraint Multiset Grammar)、PLC^[11] (Picture Layout Grammar)、LGG^[12] (Layered Graph Grammar)、RGG^[13] (Reserved Graph Grammar) 和 SCG^[14] (Spatial Graph Grammar)。

图文法用于生成和分析一个图。一个图文法通过一个初始图 and 一组产生式规则进行详细说明。产生式的作用是用一个图替换原图的一个子图。这种处理方式依赖于所需要的嵌入方法的详细说明: 在原图中连接被替换的子图的边, 怎么连接到新的被嵌入的子图。而字符文法因为是一维线性排列, 不需要特别说明这种嵌入规则, 它只需要将一个字符串插入到原字符串相应位置即可。

成员判定问题, 即文法的可归约性判断, 是指给定文法 (包含嵌入方法、产生式), 判断给定的图是否为本文法的语言。它是图文法理论研究中的关键问题之一, 用于判断一个

图是否为给定图文法产生的语言。每一种图文法形式化方法都有其自身的归约方法, LGG 的归约算法涉及两个阶段, 第一个阶段需要设计专门的算法, 而第二阶段还要对产生式的依赖关系进行判断, 从而增加了算法的时间复杂度, 其很难在实际问题应用。RGG 采用了在产生式满足 selection-free 条件时的 SFPA^[12] (Selection-Free Parsing Algorithm)。这一算法不需要进行回溯, 其时间复杂度为多项式级。但当产生式不满足此条件时, 仍将采用带回溯的算法。曾晓勤等^[15]提出了基于边的上下文相关图文法 (Edge-based context-sensitive Graph Grammar, EGG), 其主要特点是产生式的左右两端对应的图元是边, 通过边的对应关系来完成嵌入操作。EGG 中将边作为上下文元素, 更关注于图的结构信息, 在图的推导和归约过程中, 较少涉及图的语义信息, 但其归约算法仍采用带回溯的归约算法。

本文在 EGG 的形式上, 对产生式的形式进行约束, 提出了产生式形式约束条件的判断方法。通过此判断方法, 可以有效地判断产生式是否满足约束条件, 如果满足约束条件, 那么在归约过程中, 产生式的使用顺序不影响归约的结果, 从而避免了归约过程中的回溯, 可以有效提高归约的效率。

1 EGG 重要性质

定义 1 在嵌入过程中, 使用某一规则 P_i , 余图中与嵌入部分通过上下文边关联的节点, 称为外关联节点, 外关联节点集合用 VE 表示。

收稿日期: 2014-06-05; 修回日期: 2014-09-11。 基金项目: 湖北省教育厅科学研究计划项目 (Q20142607)。

作者简介: 王毅 (1980 -), 男, 湖北襄阳人, 讲师, 博士研究生, 主要研究方向: 软件可视化、图文法; 丁函 (1980 -), 女, 湖北随州人, 讲师, 硕士, 主要研究方向: 图文法、逻辑程序设计。

定理 1 在 EGG 文法中,对于任意一个产生式 P_i ,其形式为 $P_{iL} := P_{iR}, S_{iL} = \{\tilde{E}_{iL}, \tilde{E}_{iR}\}$ 和 $S_{iR} = \{\tilde{E}_{iR}, \tilde{E}_{iL}\}$ 分别表示产生式两边的悬边的集合;用 \leftrightarrow 表示悬边的对应关系,则

$$\forall E_{iL} \in S_{iL} (\exists E_{iR} \in S_{iR} \leftrightarrow E_{iL}) \wedge \\ \forall E_{iR} \in S_{iR} (\exists E_{iL} \in S_{iL} \leftrightarrow E_{iR})$$

那么,在嵌入时,有且仅有

$$\forall E_{iL} \in S_{iL} ((S(E_{iL}) \in V(P_{iL}) \wedge T(E_{iL}) \in V_E) \vee \\ (S(E_{iL}) \in V_E \wedge T(E_{iL}) \in V(P_{iL})))$$

证明 如果

$$\exists E_L \notin S_{iL} ((S(E_L) \in V(P_{iL}) \wedge T(E_L) \in V_E) \vee \\ (S(E_L) \in V_E \wedge T(E_L) \in V(P_{iL})))$$

则 $\neg \exists E_R \in S_{iR} (E_R \leftrightarrow E_L)$

而 $\forall E_{iL} \in S_{iL} (\exists E_{iR} \in S_{iR} \leftrightarrow E_{iL}) \wedge$

$$\forall E_{iR} \in S_{iR} (\exists E_{iL} \in S_{iL} \leftrightarrow E_{iR})$$

故假设不成立。所以在嵌入时,有且仅有

$$\forall E_{iL} \in S_{iL} ((S(E_{iL}) \in V(P_{iL}) \wedge T(E_{iL}) \in V_E) \vee \\ (S(E_{iL}) \in V_E \wedge T(E_{iL}) \in V(P_{iL})))$$

即在 EGG 的嵌入过程中,除了表示上下文信息的边与余图相应节点关联外,不存在其他的边与余图关联。证毕。

定理 2 在 EGG 的归约过程中,

$$\exists \text{redex1} \subset g^{\text{host}} \wedge \exists \text{redex2} \subset g^{\text{host}}$$

且 $\neg \exists V_i (V_i \in \text{redex1} \wedge V_i \in \text{redex2})$

那么: $g \xrightarrow{\text{redex1}} g' \xrightarrow{\text{redex2}} g^{*1}, g \xrightarrow{\text{redex2}} g'' \xrightarrow{\text{redex1}} g^{*2}$, 且 $g^{*1} \leftrightarrow g^{*2}$ 。

证明 因为 $\exists \text{redex1} \subset g^{\text{host}} \wedge \exists \text{redex2} \subset g^{\text{host}}$, 且 $\neg \exists V_i (V_i \in \text{redex1} \wedge V_i \in \text{redex2})$, 那么当 $g \xrightarrow{\text{redex1}} g'$ 时, $\text{redex2} \subset g'$, 从而可以得到 $g' \xrightarrow{\text{redex2}} g^{*1}$ 。同理, 当 $g \xrightarrow{\text{redex2}} g''$ 时, $\text{redex1} \subset g''$, 从而可以得到 $g'' \xrightarrow{\text{redex1}} g^{*2}$ 。故 $g \xrightarrow{\text{redex1}} g' \xrightarrow{\text{redex2}} g^{*1}, g \xrightarrow{\text{redex2}} g'' \xrightarrow{\text{redex1}} g^{*2}$, 且 $g^{*1} \leftrightarrow g^{*2}$ 。证毕。

两个图柄无共同的节点,在对主图进行归约时,两个图柄之间是相互独立的。通过对一个图柄的归约不会对另一个图柄产生影响,故对两个图柄来说,归约的先后顺序不影响归约的效果。即使两个图柄间通过表示上下文的边连接,由于在嵌入时边的对应关系,不影响使用另一个图柄进行归约。如图 1 所示。

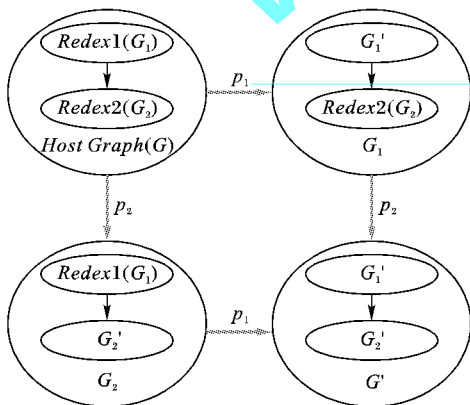


图 1 主图中两个图柄无共同节点的归约情况

2 EGG 产生式选择无关条件的判断

选择无关条件 1 (必要条件): EGG 文法中,对于任意两个

产生式 P_i, P_j , 其右部图分别为 P_{iR} 和 P_{jR} , $\neg \exists V (V \in P_{iR} \wedge V \in P_{jR})$ 。

证明 在 EGG 的归约过程中, 设 $\text{redex1} = K(P_{iR})$, $\text{redex2} = K(P_{jR})$, 且 $\neg \exists V (V \in P_{iR} \wedge V \in P_{jR})$, 所以, $\neg \exists V' \in g^{\text{host}} (V' \in \text{redex1} \wedge V' \in \text{redex2})$ 。根据定理 2, $g \xrightarrow{\text{redex1}} g' \xrightarrow{\text{redex2}} g^{*1}, g \xrightarrow{\text{redex2}} g'' \xrightarrow{\text{redex1}} g^{*2}$, 且 $g^{*1} \leftrightarrow g^{*2}$ 。

证毕。

如果两个图柄有共同的节点,如图 2(a), 要使产生式仍然选择无关, 必须保证使用产生式对其中一个图柄进行归约后, 不影响另一个图柄, 那么, 在产生式的两端须保留两个图柄共同的节点。

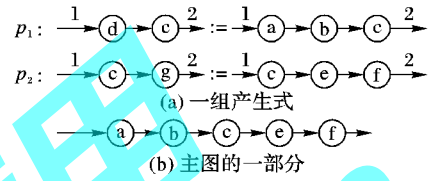


图 2 两个图柄有共同的节点

图 2(a) 的产生式两端保留了部分图元素, 为了使产生式在形式上更简单, 考虑将之改写为图 3 的产生式形式。

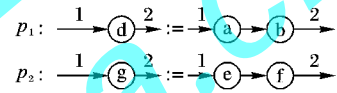


图 3 图 2(a) 改写后的产生式

通过使用图 2(a) 和图 3 两组产生式对图 2(b) 进行归约, 其结果相同。但是, 后一组产生式相对于前一组产生式, 其语法表达能力更强。

例如: 在主图中有一节点为 $\xrightarrow{1} d \xrightarrow{2}$, 通过图 3 可推导出 $\xrightarrow{1} a \xrightarrow{2}$, 但通过图 2(a) 不能推导出。而表达能力的增强也可能导致归约判断产生错误。图 4 中: L_1 为一文法产生的语言的集合, L_2 为该文法的产生式经过上述方法改写后的新文法产生的语言的集合。若需判断的图 G_1 不是 L_2 的元素, 可知其也不是 L_1 的元素, 即 G_1 不是这两个文法的语言; 但若 G_1 不是 L_1 的元素, 而是 L_2 的元素, 那么, 通过改写后的产生式可归约为其文法的语言, 但不是原文法的语言, 从而导致归约判断的错误。故此类产生式不可进行类似改写。

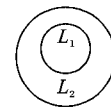


图 4 图 2(a) 的产生式改写前后文法的语言的集合

定义 2 设 $p_i \in P$, 其左部图为 P_{iL} , 右部图为 $P_{iR} \circ P_{iL}$ 的上文边集为 $E_{iL_u} = \{e \mid e \in \tilde{e} \wedge T(e) \in K(P_{iL})\}$; P_{iL} 的下文边集为 $E_{iL_d} = \{e \mid e \in \tilde{e} \wedge S(e) \in K(P_{iL})\}$ 。

定理 3 设 $p_i \in P, p_j \in P$, 其右部图分别为 $G''_{iL} = K(P_{iL}) - G'$ 和 $G''_{jL} = K(P_{jL}) - G'$ 。如果 $\exists G' (G' \subset K(P_{iR}) \wedge G' \subset K(P_{jR}))$, 那么, 当 G' 满足如下条件时, 产生式 p_i 和 p_j 是选择无关的:

- 1) $G' \subset K(P_{iL}) \wedge G' \subset K(P_{jL})$;
- 2) 若 $\exists e (s(e) \in V_E(G') \wedge T(e) \in G''_{iL})$, 那么 $e \in E_{jL_u}$;
- 3) 若 $\exists e (s(e) \in V_E(G') \wedge T(e) \in G''_{jL})$, 那么 $e \in E_{iL_d}$ 。

证明 设待归约的图为 G , 图文法中产生式组 $P, p_i, p_j \in P$, 其左部图分别为 P_{iL} 和 P_{jL} , 右部图分别为 P_{iR} 和 P_{jR} 。

$$\begin{aligned}
G &\xrightarrow{p_i} G_i \\
G_i &= (G - p_{iR}) \cup E_{iLu} \cup E_{iLd} \\
G_i &\xrightarrow{p_j} G_{ij} \\
G_{ij} &= (G_i - p_{jR}) \cup E_{jLu} \cup E_{jLd} = \\
&((G - p_{iR}) \cup E_{iLu} \cup E_{iLd} - p_{jR}) \cup E_{jLu} \cup E_{jLd}
\end{aligned}$$

而

$$\begin{aligned}
G &\xrightarrow{p_j} G_j \\
G_j &= (G - p_{jR}) \cup E_{jLu} \cup E_{jLd} \\
G_j &\xrightarrow{p_i} G_{ji} \\
G_{ji} &= (G_j - p_{iR}) \cup E_{iLu} \cup E_{iLd} = \\
&((G - p_{jR}) \cup E_{jLu} \cup E_{jLd} - p_{iR}) \cup E_{iLu} \cup E_{iLd}
\end{aligned}$$

可以得到 $G_{ij} = G_{ji}$, 故产生式 p_i 和 p_j 是选择无关的。

证毕。

选择无关条件2(必要条件): 对于一个EGG, 如果任意两个产生式是选择无关的, 那么产生式组满足选择无关条件, 其归约的时间复杂度为线性复杂度。

图5关于流程图的EGG产生式, 满足选择无关条件。从图6的归约过程可看出, 归约过程中的一个图, 如果存在多个图柄, 且图柄之间满足选择无关条件, 那么选用任意一个图柄进行归约, 不会影响归约的效果。这样, 在归约过程中没有了回溯, 使归约的时间复杂度为多项式级的。

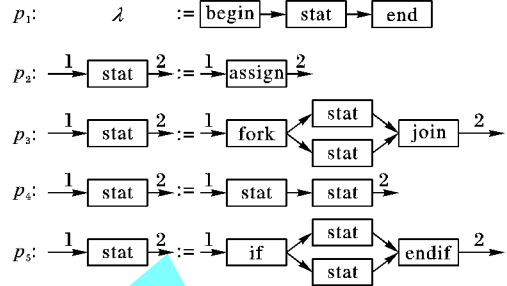


图5 一组EGG产生式

3 结语

在RGG中, 当产生式满足 selection-free 条件时, 归约过程中, 当主图中有多个图柄时, 使用图柄的先后顺序不会影响

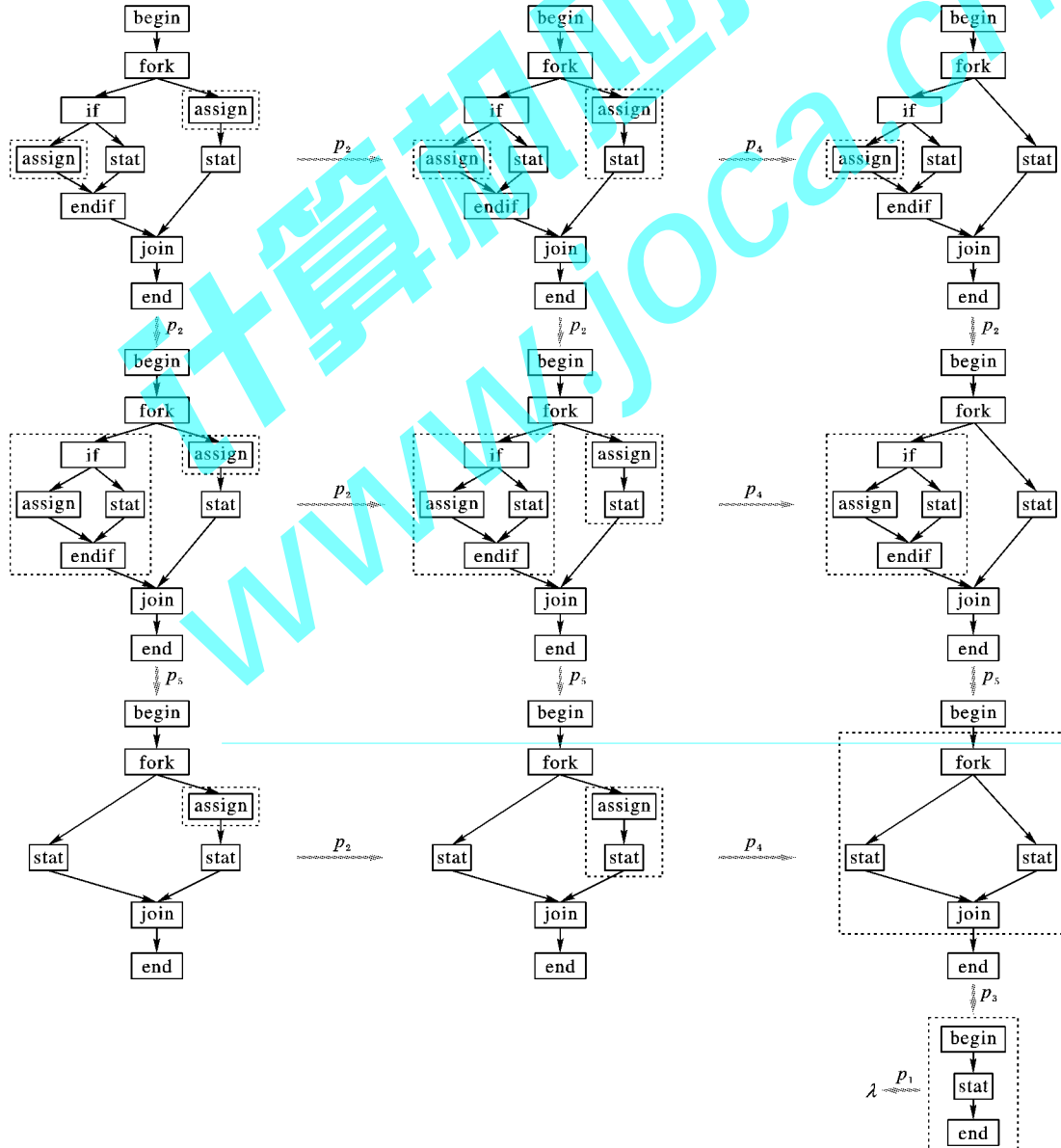


图6 一个流程图选择无关归约实例

归纳的结果。但是,由于 EGG 使用边作为上下文,在文法的形式化上和产生式的形式上都与 RGG 不同。本文根据 EGG 的这一重要特性,推导出基于 EGG 的产生式选择无关条件。当 EGG 的产生式满足选择无关条件时,归纳的时间复杂度为多项式级的。

参考文献:

- [1] PFALTZ J, ROSENFELD A. Web grammars[EB/OL]. [2010-10-10]. <http://ijcai.org/Past%20Proceedings/IJCAI-69/PDF/054.pdf>.
- [2] ECONOMAKOS G, PAPA-KONSTANTINOU G, TSANAKAS P. An attribute grammar approach to high-level automated hardware synthesis[J]. *Information and Software Technology*, 1995, 37(9): 493 – 502.
- [3] COURCELLE B. An axiomatic definition of context-free rewriting and its application to NLC graph grammars[J]. *Theoretical Computer Science*, 1987, 55(2/3): 141 – 181.
- [4] ADACHI Y, KOBAYASHI S, TSUCHIDA K, *et al.* An NCE context-sensitive graph grammar for visual design languages[C]// *Proceedings of the 1999 IEEE Symposium on Visual Languages*. Piscataway: IEEE Press, 1999: 228 – 235.
- [5] JANSSENS D, ROZENBERG G. Graph grammars with neighbourhood-controlled embedding[J]. *Theoretical Computer Science*, 1982, 21(1): 55 – 74.
- [6] Le METAYER D. Describing software architecture styles using graph grammars[J]. *IEEE Transactions on Software Engineering*, 1997, 24(7): 521 – 533.
- [7] ZHAO C, KONG J, DONG J, *et al.* Pattern based design evolution using graph transformation[J]. *Journal of Visual Languages and Computing*, 2007, 18(4): 378 – 398.
- [8] DONG J, YANG S, ZHANG K. Visualizing design patterns in their applications and compositions[J]. *IEEE Transactions on Software Engineering*, 2007, 33(7): 433 – 453.
- [9] FLASINSKI M. Power properties of NLC graph grammars with a polynomial membership problem[J]. *Theoretical Computer Science*, 1998, 201(1/2): 189 – 231.
- [10] MARRIOTT K. Constraint multiset grammar[C]// *Proceedings of the 1994 IEEE Symposium on Visual Languages*. Piscataway: IEEE Press, 1994: 118 – 125.
- [11] GOLIN E J. A method for the specification and parsing of visual language[D]. Providence: Brown University, 1991.
- [12] REKERS J, SCHÜRR A. Defining and parsing visual languages with layered graph grammars[J]. *Journal of Visual Languages and Computing*, 1997, 8(1): 27 – 55.
- [13] ZHANG D, ZHANG K, CAO J. A context-sensitive graph grammar formalism for the specification of visual languages[J]. *The Computer Journal*, 2001, 44(3): 186 – 200.
- [14] KONG J, ZHANG K, ZENG X. Spatial graph grammars for graphical user interfaces[J]. *ACM Transactions on Computer-Human Interaction*, 2006, 13(2): 268 – 307.
- [15] ZENG X, HAN X, ZOU Y. An edge-based context-sensitive graph grammar formalism[J]. *Journal of Software*, 2008, 19(8): 1893 – 1901. (曾晓勤, 韩秀清, 邹阳. 一种基于边的上下文相关图文法形式化框架[J]. *软件学报*, 2008, 19(8): 1893 – 1901.)

(上接第 3179 页)

参考文献:

- [1] ADLEMAN L M. Molecular computing of solutions to combinatorial problems[J]. *Science*, 1994, 266(5187): 1021 – 1024.
- [2] GAO T, CHEN Z. Image encryption based on a new total shuffling algorithm[J]. *Chaos, Solitons and Fractals*, 2008, 38(1): 213 – 220.
- [3] ARROYO D, LI C Q, LI S J, *et al.* Cryptanalysis of an image encryption scheme based on a new total shuffling algorithm[J]. *Chaos, Solitons and Fractals*, 2009, 41(5): 2613 – 2616.
- [4] NING K. A pseudo DNA cryptography method[EB/OL]. [2012-10-10]. <http://arxiv.org/pdf/0903.2693v1.pdf>.
- [5] XUE X. Research on digital image encryption based on DNA sequence and multi-chaotic maps[D]. Dalian: Dalian University, 2010. (薛香莲. 基于 DNA 序列与多混沌映射的数字图像加密技术研究[D]. 大连: 大连大学, 2010.)
- [6] ZHANG Q, ZHOU S, WEI X. An efficient approach for DNA fractal-based image encryption[J]. *Applied Mathematics and Information Sciences*, 2011, 5(3): 445 – 459.
- [7] ZHANG X, XU G, FU X. Research on spread spectrum watermarking algorithm resisting image-cropping based on baker's transformation[J]. *Application Research of Computers*, 2012, 29(6): 2246 – 2248. (张鑫, 徐光宪, 付晓. 基于面包师变换的抗剪切扩频水印算法研究[J]. *计算机应用研究*, 2012, 29(6): 2246 – 2248.)
- [8] LIU L, ZHANG X. Image encryption algorithm based on chaos and bit operations[J]. *Journal of Computer Applications*, 2013, 33(4): 1070 – 1073, 1099. (刘乐鹏, 张雪峰. 基于混沌和位运算的图像加密算法[J]. *计算机应用*, 2013, 33(4): 1070 – 1073, 1099.)
- [9] JIANG J, YIN Z. The advantages and disadvantages of DNA password in the contrast to the traditional cryptography and quantum cryptography[J]. *Science and Technology Vision*, 2012(24): 24 – 27. (蒋君, 殷志祥. DNA 密码对比传统密码学与量子密码学的优势与不足[J]. *科技视界*, 2012(24): 24 – 27.)
- [10] RAHIMOV H, BABAEI M, HASSANABADI H. Improving middle square method RNG using chaotic map[J]. *Applied Mathematics*, 2011, 2(4): 137 – 141.
- [11] SADEG S, GOUGACHE M, MANSOURI N, *et al.* An encryption algorithm inspired from DNA[C]// *Proceedings of the 2010 International Conference on Machine and Web Intelligence*. Piscataway: IEEE Press, 2010: 344 – 349.
- [12] BAO G, JI S, SHEN J. Magic cube transformation and its application in digital image encryption[J]. *Journal of Computer Applications*, 2002, 22(11): 23 – 25. (鲍官军, 计鸣, 沈建冰. 魔方变换及其在数字图像加密中的应用[J]. *计算机应用*, 2002, 22(11): 23 – 25.)
- [13] HAN F, ZHU C, HU Y. New colour image encryption algorithm based on high-dimension chaotic system[J]. *Journal of Computer Applications*, 2007, 27(8): 1888 – 1890. (韩凤英, 朱从旭, 胡玉平. 一种基于高维混沌系统的彩色图像加密新算法[J]. *计算机应用*, 2007, 27(8): 1888 – 1890.)
- [14] HUANG J. Research on the secure chaos-based image encryption algorithm[D]. Changsha: Changsha University of Science and Technology, 2010. (黄金. 安全混沌图像加密算法的研究[D]. 长沙: 长沙理工大学, 2010.)