

RECO - ÍÐÎÃÐÀÌÍÀß ÑÈÑÒÅÌÀ ÄËß ÐÀÑÏÇÍÀÅÀÍÈß ÑÒÀÐÎÑËÀÂßÍÑÊÈÕ ÒÅÊÑÒÔÂ

Àâòîð Àëåêñàíäð Âàëåðûââè÷ Êó÷óâàíâ

09.07.2010 á.

Íñëåäíâ íáñâäéâíéâ 15.07.2010 á.

In the paper software system RECO for the recognition of Old Slavonic texts is described. Methods and technology are described for the analysis of the vector patterns of the letters. Examples of Old Slavonic text recognition are given.

Iðè

ñíçääìèè ýëåêòðííûõ êïëëåêöèé ñòàðñâàòíûõ òåêñòâi (ðóëïëñè 15 – 18 ååêîâ) ñâðüâçííé iðíáëâííé ïñòàþòñý çàäà÷è ìáðàáîòêè è ðàñïïçíàâàíèÿ èçïáðàæåíèé. Å óñëâèÿð, êïäàà iðèäëíàëû àïëôíàíòâ ñèëüíí èçïíøâíû, öââò áóïàäè íâíäíðíäáí, èìââðñý iííæåñòâi ïòåðòïñòâé, çàðâííàéé, iýoâí, áïëåå iíçäíèõ ïílåòiè, iðåäúÿâëýþòñý åúññèå òðåáâàíèÿ è iðíäðàìà ìáðàáîòêè èçïáðàæåíèé è ðàñïïçíàâàíèÿ [Êíííüþòåðíâ 2006].

Iieñüâàâàÿ

íèæå iðíäðàìàÿ ñèñòâà Reco – ýëñïåðèìâòàëüíàÿ ååðñèÿ, iñíáâííñòÿè åå ýâëýþòñý:

1)

èñïïëüçîâàíèå ñòðóêòóðíïäî
(ëèíâèñòè÷åñêîñî) ïïäõñää è ðàñïïçíàâàíèþ áóêå;

2)

ðàçâèòûé ðåæèì ïðåäíáðàáîòêè,
ðåâàéòèðîâàíèÿ è ïïïíëíåíèÿ ýòàëíïâ.

Íðîñåðàììà

ñîñáðæèò ñòàíäàðòíûé íàáîð èíñòðóìåíòâ äëÿ ïïâûðåíèÿ êà÷åñòâà èçîáðàæåíèÿ:
ñâéàæèâàíèÿ, ôèëüòðàöèè, êíòðàñòèðîâàíèÿ, ïïñòðóìåí÷àòîé ñòàòèñòè÷åñêîé
áéíàðèçàöèè.

Ýòàï

íáðàáîòêè – âûäåëåíèå ñêåëåòíïà.

Ðèñ. 1. Ðåêóðñèâíïå óâåëè÷åíèå

ñòåïåíè äåòàëèçàöèè

Iðeíöeí

Iñéó÷áíèý ñêåëåòííà çàéëþ÷áòñý á ðàçáèåíèè ôèäóðû íà áúñóéëñá iáëæñòè á òåð
ìñòáð, áäá áðàíèöà èíååò ìðèòåòåëüíûé iáðåáéá [Ðàñííçíáâàíèå 2008, n. 168-172].

Iðe ýóíí ðåéóðñèáíí óíáíüøååòñý áåééè÷éíà iòñåéàåíûó iáëæñòåé, ò.å.

Óåééè÷éååòñý ñòåííáü ååòåéçåöèè áíáééçéðóåííáü iáúåéòà (ðèñ. 1, 2).

Ðèñ. 2. Äåóõñòóíáí÷àòàÿ
áéíàðèçåöèý è ñêåëåòíí

Iðåäíáðåáíòéá

Ñêåëåòííáí òçíáðåæåíèý. Á iðåäíáðåáíòéó iñíèíí áíðíéñèíèðíåáíèý áóñáííáí
ñêåëåòííáí òçíáðåæåíèý áéëþ÷áíû iñíðåòéè iñíðåííèý ááí áðåôà è iñðåäåéåíèý
éà-áñòååííûó iñíðåíåðåíáü iáúåéòíá òçíáðåæåíèý (ðèñ. 3).

Ðèñ.

3. Ðåæèí iðåäíáðåáíòéè

Ðåáíòá

Ñ ýòåéííáí. Á iñíðåðåííá èíååòñý áíçííæíñòü iñíñíàòðéååòü, ðåäåéòðéååòü, äíáàåéýòü

è óääëëýöü ýòàëëííû.

Äëý

äíáàâëåíèý ýòàëëííá íåíáñíäèíí áúñíëíèòü íðåäáñòááòéó ñêåëåòííáí èçíáðàæåíèý, áúáñòáòü áóéâó è íàæàòü íá éíííéó «Níñòáíèòü éàé ýòàëëí ->». Iíjyäéòñý íéíí ááíäà èíäíé ýòàëëííá (ðèñ. 4). Á äàíííí íéíá íåíáñíäèíí áúáñòáòü ñðèôò, á éíòíðíí áóáñòü áúáñäëòüñý èíý. Eíý ááíäëòñý á ñíñòááòñòáóþùáí ííëá èëë áúáëðàáñý èç òàäëëöü ñèíäíííá áúáñòáííñí ñðèôòòá. Iíñëá ááíäà èíäíé íåíáñíäèíí íàæàòü íá éíííéó «Ok». Iðíñòáííá äíáàâëòü ííáñé ýòàëëí á éííñöü ñíèñéà. Áñëè ííä áóéâíé íáñòáëëñü ííëäíæéà, óí è ýòàëëíí ñðèëíæëòñý è ááí èññòáííá èçíáðàæåíèá.

Đèñ.

4. Äíáàâëåíèå ýòàëëííá

Èíååòñý

âíçíííñòü íáçíá-àòü ááñ äëý óçëíá è áåòâåé – ñòáíííü èõ áàæííñòé, ó-èòúâåáíàÿ íðè ñíñòáâëåíèé ñòàòèñòéèé ñíáíáäåíèý íðè ðàñííçíáâåíèé. xòíáú íáçíá-èòü ááñ, íåíáñíäèíí áúáñòáòü óçáé èëë áåòâü, áâåñòé çíá-åíèá á ííëá «Áñ» è ñíýòü áúáåëåíèá ñ ýéáíåíòá.

Çíá-åíèý áåñíá ííäóó áúòü íðíèçâíëüíùé, áàæíí èëëü èõ áçàëííá ðàçëè-éå.

Ýòàëëííü ñòáíáðàæàþòñý ñ ó-åòíí áåñà óçëíá è áåòâåé: íáíáå áåñííúå ýéáíåíòü íòíáðàæàþòñý óííüøá è ñâåòëåå (ðèñ. 5).

Đèñ.

5. Äèçóàëëçàöëý ýòàëëííá ñ ó-åòíí áåñà óçëíá è áåòâåé

Ðàñííçíáâåíèå.

Ðåçóëüòàò ðàñííçíáâåíèý íáðåäåáñòñý á òåéñòíáûé ðåäääéòíð Microsoft Word (èññíëüçíáàíû ñðèôòòü ñ ñàéòá: <http://www.irmologion.ru/>).

Đèñ.

6

Íá áàííúé

ííáíò íðíñòáííá Reco äàåò 86% ðàñííçíáííúó ñèíáíëíá íá ðóéñëñýó ñíðíøåáí êà-åñòáà è 64% íá ðóéñëñýó íéçéíáí êà-åñòáà.

EEÒÅÐÀÓÐÀ

1. Êó÷óãàíâ Å.Â., ïñêîëêîâ I.I.

Đàñïïçíàâáíèå ñòàðîñëàâýíñèø òåëñòâi ìåòîäàìè, ïñííâàííùè íà áèîàëäîðèòìàõ àíàëèçà èçíàðàæåíèé. // Ôðóäû lâæäöíàð. íàó÷. êíîô. (Êàçàíü, 26-30 àâãñòà 2008 å.). – Èàçàíü: Èçä-âî Èàçàí. åññ. óí-òà, 2008. – N. 168-172.

2. Øäïëðî

Ë., Ñòîëìàí Äæ. Êññüþòåðíâ çðåíèå. Íåð. ñ àíãë. – I.: ÁÈÍÌ, 2006.