

Âúââèæáíèà è ìðíâðèà àèìòàç â ñèñòàíà ðàñíïçíàâàíèÿ äðâáíáðóññèé ñèìðíèèè

Ààðìð Èàáí Àíàðìèùààè- Çàèáíòíà
10.07.2010 ä.
İñèàáíáà íáííàèáíèà 10.07.2010 ä.

Electronic processing of ancient Russian text requires its representation in electronic text format. An OCR system may help in translation of real documents into electronic text. A structure-based hypothesis-driven recognition method is proposed. The paper describes the principles of hypotheses nomination and testing used by this method.

Äèÿ ððàíáíèÿ, ÿèàèððííííàí èççààíèÿ è èíííÿððàððííè íáðâáíðèè ìàìÿòíèèíà äðâáíáðóññèé ñèìðíèèè ððâáóòòíÿ èð ìðââñòàèèà çíáíèè. Ýòí ìáñòíÿòáèÿñòàí ààèààð ðàèàñííàðçàçíí ñíçààíèà ààðííàðèçèððíáííè èíííÿððàððííè ñèñòàíú ðàñíïçíàâàíèÿ àíèòíáíòíà ÿò ðàñíïçíàâàíèÿ, ìñòðíáííàÿ ìà ñèàáòÿèð èð ðàðàèððàððííè ìðèèèèàð: ñòðòèèòòðííèè ìððíà è ðàñíïçíàâàíèèð [Ñòðòèèòòðííèè 1977], ððàñí àáòòððíáííàÿ ìðâáíèçàòèèÿ ìðíòàññà ðàñíïçíàâàíèÿ (ñèíáà è áóèâú), ìá-, òèíà ìèèñàíèà ñòðòèèòòðííèè ÿèàíáíòíà ðàñíïçíàâàíèèò ìá ìðââèàâàòòíÿ èñèèð-àòú ìóð, ì òíðàèèáíèÿ ìðíòàññíí ðàñíïçíàâàíèÿ àúââèæáíèàí àèìòàç ì ðàñíïçíàâàíèèò ìáúàèòàð è èð ìðíáðð ìáðçíí, àúââèÿðòíÿ òíèèèí ñòúáñòàáííúà -àñòè èçíáðâæáíèÿ, à ìñòíðííèèà ìñòàðòíÿ áàç áíèìáíèÿ. Ðàçðâáíòáí ñííííà ìèèñàíèÿ ñ 1979]. Ñòðòèèòòðà ñèíà è áóèà èàðððèè-àñèè ìáðçíí ìðââñòàèèÿòòíÿ à áàçà çíáíèè ñèñòàíú. Ìñííáíúà òçèú ððàèíáíè ñàòè èí ìáðâñá-áíèÿ. Ðàñíïçíàâàíèà áóèà è ñèíà ñòðíèèòíÿ ìá ìñííàà èííòáíòèè àèðòòàèèííáí òðàèíà. Äèÿ ìðââèè, ìííòèè ðàññííòðèè ìðíòàâ ìèñòààðÿàè ìáàèðáàáíòð èàðòèíò. Ýòíò ððàèí ìàçúâààòòíÿ àèðòòàèèííú. Çààà-à ðàñíïçíàâàíèÿ ñáíàèòíÿ è ìáðíæááíèèð ñííííàà ñòðòèèòòðà àèðòòàèèííáí òðàèíà ìçáíèÿò àúââèèòú ìááíð òðàèíá áàçú çíáíèè à èà-àñòàà ñèèèà àèìòàç, ìðáíòèèèè ìèñòàâ ññòèíè ìá ñíæàñíáííúà òçèú è òáí ñáíú ìèñòààðÿàÿ ñòáíò ñíæàñíáíèÿ. Äèÿ ìðââèèáíèÿ ìðââñííáíííòè àèìòàç èñíèèçòòò ìðââñèàâàííò òðàèíò è ìðââèèÿòòíÿ èàè ìòííáíèà -èñèà ìáð ñíæàñíáííúò òçèíà à àèìòàçà è ìáúáíò -èñèò òçèíà à òðàèíà á àèìòàçú àíáíðèò ì òí-ìíòèè òñòàííàèáííáí ñíòààòòòàèÿ è àú-èñèÿòòíÿ èàè ìòííáíèà -èñèà ìáð à àèìòàçà è -èñèò òçèíà ÁÔ. àèìòàçú. Ìðè ìðíáððèà àèìòàç àèÿ èàæáíè èç ìèð àú-èñèÿòòíÿ ñòáíáíè ñíæàñíáíííòè è ìðèáíííòè è ìáíà èç ìèð ìðèíèààòñ èèíèè. Ààèàà ððàññèððíáúèèò àðíáííáí èçíáðâæáíèÿ ìðââààðòíÿ çáíðííú ìá ìèèè ìæèàááíúò èèíèè ìá èçíáðâæáíèè, ìèò-áííúà ðð òðàèíàè ààçú çíáíèè àí àñàò àèìòàçàð. Ààèàà àúñíèÿòòíÿ ìáðâñ-, ò ðàðàèððàððííèè àèìòàç è ìáðòèèèèèà òñèíàèà ìðèáíííòè òèàçúâàáíàÿ àáííè àèìòàçíè. Áñèè òàèíè àèìòàçú ìáð, òàèóúàè ìàçíà-ààòíÿ àèìòàçà ñ ìàèñèíàèèííè ñòáíáíúð ñíæàñíáíííòè ñíæàñíáíèà òçèíà áóèà è ìðíòðàíñòàáííúò ìòííáíèè ìæàò ìèèè. Èèòàðàòòðà1. xáðáííèí È.Á. Ðóññèàÿ ìàèáíáðàòèÿ Ì., 1956 ðàñòðíáúò ìííòðíííúò èçíáðâæáíèè àèàðèáííáí ðââàèòíðà SpotLight // Èíòàèèàèòòàèèííúà òàðííèíàèè è ñèñòàíú: ñáíðíèè ñòàòàè àèáíðèòíú àíàèèçà èííòòðííúò èçíáðâæáíèè à àèçòàèèííúò èííòíàòèèííúò ñèñòàíàð ìá ìñííàà ìáíáííðíáííè ìá-, òèíè ñáíáíòè-àñèíè ñá 2004.5. ìèíèè Ì. Òðàèíú àèÿ ìðââñòàèèáíèÿ çíáíèè. Ìáð. ñ áíàè. 1979.