

**Д.Г. ШАПОШНИКОВ, Г.Г. БОНДАРЬ, В.И. ГУСАКОВА,
Л.Н. ПОДЛАДЧИКОВА**

**НИИ нейрокибернетики им. А.Б. Когана Ростовского государственного
университета
nisms@krinc.ru**

РАСПОЗНАВАНИЕ ДОРОЖНЫХ ЗНАКОВ: МОДЕЛЬ И ЭКСПЕРИМЕНТ

Аннотация.

Исследование проведено на 20 изображениях дорожных знаков (ДЗ), сгруппированных попарно по максимальной близости признаков описаний в модели FOSTS [1-4]. В психофизических тестах, имитировавших различные ситуации дорожного движения, оценивалась вероятность ошибок при восприятии изображений выбранных ДЗ. По результатам психофизических и модельных тестов выявлены ДЗ, вызывающие максимальное количество ошибок при распознавании. Рассматриваются возможные пути увеличения эффективности распознавания ДЗ в модели и модификации изображений ДЗ, «сложных» для зрительного восприятия в критических дорожных ситуациях.

1. Введение

В связи с возрастающими в последние годы проблемами безопасности дорожного движения все большее внимание уделяется разработке интеллектуальных средств поддержки водителей, существенными компонентами которых являются системы распознавания дорожных знаков [5-9]. Актуальная задача в этой области состоит в поиске возможных путей повышения эффективности систем распознавания ДЗ различного типа (а не только ограничений скорости) в изменяющихся условиях реального дорожного движения [7].

В данной работе рассматривается подход к поиску путей увеличения эффективности распознавания ДЗ с помощью системы FOSTS, разработанной ранее [1]. Он состоит в детальном сопоставлении результатов модельных и психофизических тестов при распознавании изображений одних и тех же ДЗ. Представлены первые результаты реализации этого подхода.

2. Модель распознавания дорожных знаков FOSTS

Базовые процедуры. Ранее была разработана фовеальная модель распознавания дорожных знаков (Foveal System for Traffic Signs recognition, FOSTS, [1]). Она основана на имитации ключевых механизмов биологического зрения [1-4], таких как пространственно-неоднородное (фовеальное) представительство изображений ДЗ, ориентационная избирательность нейронов, контекстное описание элементов изображений, детальная обработка информативных фрагментов изображений и особенностей цветового восприятия человека. Модель состоит из нескольких модулей, выполняющих следующие процедуры: цветовая сегментация; классификация ДЗ по цвету и форме внешнего контура; выделение «чистого» ДЗ; представление ДЗ как n -мерного признакового вектора с помощью пространственно-неоднородного входного окна, расположенного в центре ДЗ и распознавание с помощью сравнения признаковых векторов текущего изображения и изображений прототипов ДЗ. В соответствии с задачами данного исследования были использованы два модуля модели, операции в которых коротко рассмотрены ниже.

Признаковое описание ДЗ формировалось с помощью пространственно-неоднородного сенсора (входного окна), содержащего один базовый (центральный) узел и 48 периферических узлов, расположенных на пересечении 3-х концентрических окружностей и 16-ти лучей, исходящих из центра (Рис. 1). Ориентация перепадов яркости в окрестностях каждого из 49 узлов сенсора определялась с помощью вычисления разности между двумя ориентированными Гауссианами с пространственно сдвинутыми центрами. Пространственно-неоднородное представление изображений ДЗ эмулировалось с помощью свертки Гаусса, размер ядра которой увеличивался от центра входного окна к его периферии. Размер входного окна – 36 пикселей, размер ядра в свертке Гаусса в центральной, промежуточной и периферической частях входного окна был равен, соответственно: 5x5, 7x7 и 9x9 пикселей. Кроме того, для детализованного описания ДЗ оценивалось распределение ориентированных элементов в контекстных областях 48 узлов входного окна (за исключением центрального узла). Такая структура входного окна и его расположение в центре ДЗ обеспечивало максимальное представительство ориентированных элементов (до 90%) внутренней информативной части ДЗ на первом и втором уровнях разрешения.

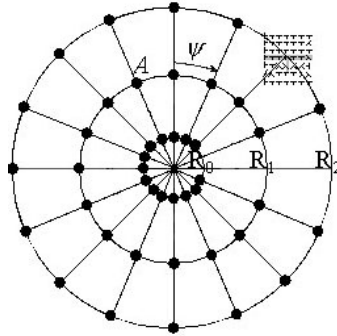


Рис.1. Структура пространственно-неоднородного сенсора. В виде решетки показана контекстная область одного из периферических узлов.

Распознавание изображений дорожных знаков проводилось по отдельным подгруппам (по цвету и форме внешнего контура) с помощью сопоставления признаковых векторов, описывающих текущее изображение, с хранящимися в базе векторами-эталонами, описывающими стандартные дорожные знаки. Сравнение признаковых векторов проводилось по формуле (1). Изображение считалось распознанным, если выполнялись два условия – значение K^b должно быть больше 25 и разность между максимальным значением коэффициента близости K^b и ближайшими к нему по величине должна быть более 4 (параметры определены экспериментально).

Коэффициент близости K^b определялся следующим образом:

$$K^b = \sum_{i=0}^{48} \left[Sg(OR_i^b - OR_i^{rw}) \cdot (1 - abs(\rho_i^b - \rho_i^{rw})) \right] \quad (1)$$

где $Sg(x) = \begin{cases} 1, & \text{если } x = 0; \\ 0, & \text{в противном случае} \end{cases}$

OR_i - доминирующая ориентация перепада яркости в контекстной области данного узла сенсора; индекс b присвоен изображениям из данной подгруппы базы эталонов, индекс rw – изображению, предъявляемому для распознавания; ρ - доля доминирующего перепада яркости в контекстной области.

Тестирование модели. Для оценки эффективности распознавания ДЗ различного типа в модели и ошибок зрительного восприятия человека в психофизических тестах были выбраны по 10 запрещающих (с внешним контуром треугольной формы и красного цвета) и 10 предупреждающих

(с внешним контуром круглой формы). Для проведения психофизических тестов (см. ниже) эти изображения были сгруппированы попарно по максимальной близости признаков векторов, описывающих эти ДЗ в модели (Рис. 2).



Рис. 2. Изображения 20 дорожных знаков, выбранных для модельных и психофизических тестов. Знаки, сгруппированные в пары ($n=10$), близкие по признаковым описаниям внутренней части изображений, расположены друг под другом.

Для тестирования модели FOSTS ранее [1] была использована вся база имеющихся ДЗ, как стандартных, так и полученных в условиях реального дорожного движения (вероятность распознавания в последнем случае составила $p=0,97$). В условиях искусственных трансформаций (зашумление, масштабирование, перспективные искажения) изображений всей выборки стандартных ДЗ обнаружено более выраженное снижение эффективности распознавания запрещающих ДЗ по сравнению с предупреждающими [1,3]. Для частной выборки, использованной в данной работе и представленной на Рис. 2, каждый ДЗ был зашумлен (50% гауссовский шум) и последовательно уменьшен от 60% до 20% своего начального размера с шагом в 5%. Затем была определена вероятность распознавания каждого знака при разных масштабах. По результатам этих тестов различий между группами запрещающих и предупреждающих ДЗ не обнаружено ($p=0,61$ и $p=0,71$). При этом как устойчивые, так и чувствительные к трансформациям изображения ДЗ представлены в обеих группах (Рис. 3).

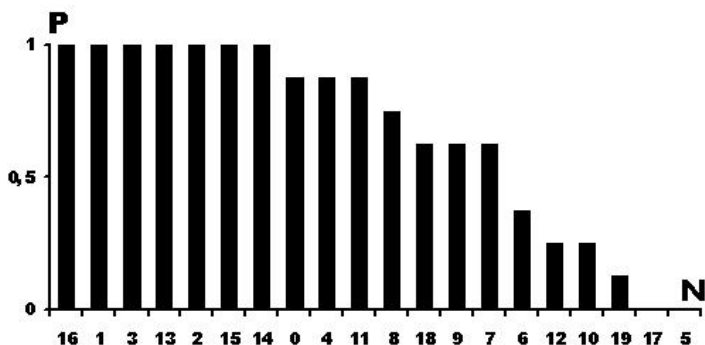


Рис.3 Результаты распознавания трансформированных дорожных знаков в модели.

3. Результаты психофизических исследований.

Те же изображения 20-ти ДЗ, которые применялись для тестирования модели, использованы при проведении психофизических тестов, целью которых была оценка точности распознавания различных ДЗ. В отличие от большинства аналогичных работ [10-12], в которых оценивались понятийные качества различных ДЗ, данное исследование направлено на тестирование перцептуальных качеств ДЗ на уровне, в основном, преаттентивных механизмов. С этой целью использовалось кратковременное предъявление (50 мс) изображений ДЗ на экране дисплея. Было поставлено несколько серий экспериментов, имитировавших различные ситуации дорожного движения, результаты которых сходны на качественном уровне у всех испытуемых. В данной работе представлены результаты двух серий тестов, в первой из которых использовались незашумленные изображения ДЗ, во второй – изображения, на которые был наложен 50% гауссовский шум.

Угловой размер изображений, предъявляемых испытуемым, соответствовал расстоянию 80 м до знака. Изображения ДЗ экспонировались на дисплее последовательно с интервалом 70 мсек между стимулами, подаваемыми попарно, время между предъявлением последовательных пар в тесте составляло 3 с. При тестировании пары стимулов составлялись: (а) из двух одинаковых ДЗ; (б) из двух похожих ДЗ, при этом очередность первого и второго стимулов в каждой паре и последовательность пар изменялись в каждом тесте случайно, всего в каждом тесте предъявлялось 40 пар ДЗ и он занимал не более 130 сек. Испытуемым предлагалось с помощью

нажатия клавиши «мыши» максимально быстро отдифференцировать ситуации типа (а) от ситуаций типа (б).

Основная часть тестов проведена на 10 студентах (средний возраст – 21,5 года). Нарушения рефракции (при их обнаружении у испытуемых) корректировались с помощью подбора линз. По результатам повторных тестов (n=6) на одном и том же испытуемом определялась вероятность правильного ответа и время реакции (латентный период между выключением второго стимула и нажатием на клавишу манипулятора «мышь»).

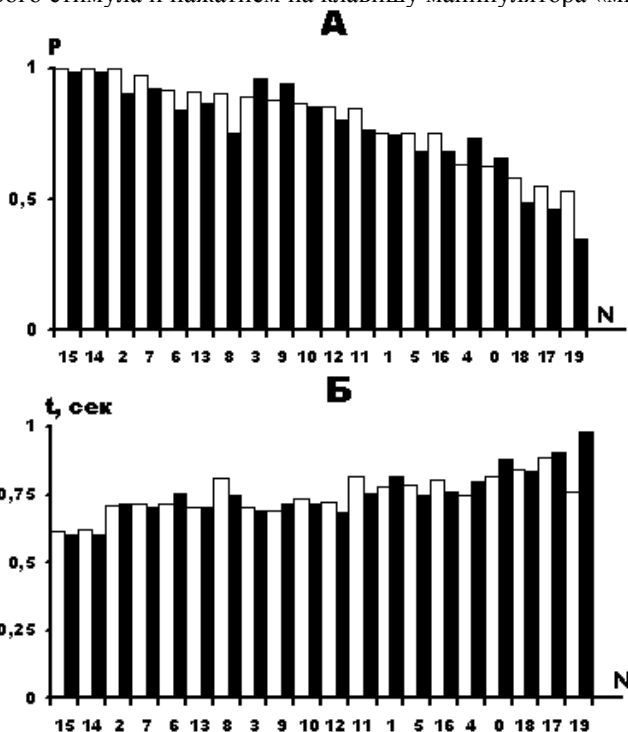


Рис. 4. Результаты психофизических тестов. А – средняя вероятность правильного ответа испытуемых при предъявлении изображений различных ДЗ; Б – средний латентный период ответов испытуемых (от 600 мсек до 970 мсек); белые столбики – тесты с предъявлением незашумленных стимулов; черные столбики – тесты с зашумленными изображениями; все диаграммы упорядочены по оси X (номера изображений ДЗ) в соответствии с вероятностью правильного ответа на незашумленные ДЗ.

Результаты, полученные в двух сериях психофизических экспериментов, представлены на Рис. 4. Видно (Рис. 4, А), что при среднем снижении вероятности безошибочного восприятия зашумленных изображений ($r=0,81$ и $r=0,77$ соответственно) различные ДЗ имеют сходный тренд распределения вероятностей в двух сериях экспериментов ($r=0,92$, $p<0,01$). Аналогичное сходство распределений заметно и по латентному периоду ответов (Рис. 4, Б, $r= 0,77$, $p<0,01$). При этом отчетливо прослеживается тренд увеличения латентного периода от ДЗ с высокой вероятностью правильных ответов к знакам с низкой вероятностью ($r= 0,75$, $p<0,01$).

4. Заключение.

При тестировании одних и тех же изображений ДЗ на модели FOSTS и в психофизических экспериментах обнаружено, что «сложные» для распознавания знаки имеются как в группе предупреждающих, так и в группе запрещающих ДЗ. Если первая группа ДЗ менее критична для безопасности дорожного движения, то низкий уровень корректности распознавания некоторых запрещающих ДЗ (например, № 5 на Рис. 2), как в модели, так и в эксперименте, ставит вопрос о необходимости оптимизации изображений этих ДЗ в соответствии с принятыми стандартами [13]. Такая оптимизация позволит снизить количество зрительных ошибок водителей в критических дорожных ситуациях.

В контексте возможного практического использования модели FOSTS как средства поддержки водителей следует отметить существенное различие эффективности распознавания большинства ДЗ в модели и в эксперименте (ср. Рис.3 и Рис. 4, $r= -0,52$, $p<0,05$). Одна из возможных причин различий результатов модельного и экспериментального тестирования может быть связана с тем, что эффективность распознавания в первом случае определяется, в основном наличием мелких деталей на изображениях, во втором – мерой близости изображений в парах ДЗ. Для приближения модели к эффективности распознавания ДЗ человеком необходим ряд ее модификаций. В частности, должны быть модифицированы некоторые процедуры и параметры входного окна модели для более корректного признакового описания, а также разработаны адаптивные алгоритмы оценки необходимой и достаточной детализации описаний конкретные ДЗ.

Работа поддержана РФФИ, грант № 05-01-00689.

1. Gao X., Shaposhnikov D., Gusakova V., Hong K., Podladchikova L. The foveal system for traffic signs recognition (FOSTS). In Proc. XIV Int. Conf. on Neurocybernetics, Rostov-on-Don, (2005), V.2, 223-227.
2. Шапошников Д.Г., Головань А.В., Подладчикова Л.Н., Шевцова Н.А., Гусакова В.И., Гизатдинова Ю.Ф. Применение поведенческой модели зрения ВМВ для инвариантного распознавания лиц и дорожных знаков. Нейрокомпьютеры: разработка и применение (2002), №7-8, 21-33.
3. Gao X.W., Podladchikova L., Shaposhnikov D., Hong K., Shevtsova N. Recognition of traffic signs based on their colour and shape features extracted using human vision models. J. Vis. Commun. Image R. 17, (2006), 675-685.
4. Rybak I., Gusakova V., Golovan A., Podladchikova L., Shevtsova N. Attention-guided recognition based on "What"; and "Where" representations: A behavioral model. Neurobiology of Attention, Elsevier, Academic, (2005), 663-670.
5. Gavrila D. M., Franke U., Gorzig S. and Wohler C. Real-time Vision for Intelligent Vehicles. In Proc. IEEE Instrumentation and Measurement Magazine. (2001), 4. 22-27.
6. Hsien J.-C. and Chen S.-Y. Road Sign Detection and Recognition Using Markov Model. 14th Workshop on OOTA, (2003), 529-536.
7. Miura J., Kanda T., and Shirai Y. An Active Vision System for Real-Time Traffic Sign Recognition. In Proc. IEEE Int. Conf. on Intelligent Transportation Syst. (2000), 52-57.
8. Shneier M. Road Sign Detection and Recognition. Subm. to the IEEE Com. Soc. Int. Conf. on Comp. Vis. and Pat. Rec., (2005). http://www.isd.mel.nist.gov/documents/shneier/Road_Sign_Detecton.pdf
9. Yang H.-M., Liu C.-L., Liu K.-H., and Huang S.-M. Traffic sign recognition in disturbing environments. Lect. N^o. in Comp. Science, V. 2871, (2003), 252-261.
10. Al-Madani H. M. N., Al-Janahi A. R. Warning signs mostly associated with drivers accident involvement. Proc. The 3rd Int. Conf. on Traffic & Transport Psychology. (2004), www.ictp.com/.
11. Waard D., Mesken J., and Brookhuis K. A.. Recognition of electronic road signs. www.hfes-europe.org/books/firstpage1999/6.pdf
12. Zakowska L. Perception and recognition of traffic signs in relation to drivers characteristics and safety – a case study in Poland. 14th ICTCT workshop proceedings (2000) <http://www.ictc.org/workshops/01-Caserta/Zakoska.pdf>
13. The improvement of road signing in Europe, ERF Road Safety Programme/, 2nd European Road Congress, (2003), <http://www.erf.be/content/general/dtair/2491>