

Автоматизация ввода паспортных данных студентов в систему управления

Ковалева И.Л., Суворов А.А., Кирцун В.М.

Белорусский национальный технический университет

Автоматизированный ввод паспортных данных – один из вариантов заполнения базы данных, содержащих личные данные студентов. Автоматизированный ввод подразумевает предварительную обработку сканированного изображения страницы паспорта и распознавание найденной информации.

Для предварительной обработки используются типовые алгоритмы перевода цветного изображения в полутоновое (в случае цветного сканирования) и бинаризации изображения. Затем выполняется сегментация. Для сегментации области, содержащей личные данные студентов, используется эталон, содержащий изображение букв «ПР», являющихся начальными буквами слова «ПРОЗВІШЧА» на 33-й странице паспорта гражданина Республики Беларусь. Область, находящаяся под словом «ПРОЗВІШЧА» передается для дальнейшей обработки. Сегментация букв в выделенной области начинается с сегментации строки. Алгоритм сегментации строк основывается на том, что средняя яркость в изображениях межстрочных промежутков существенно ниже средней яркости в изображениях текстовых строк. Следующим шагом является сегментация букв, которая схожа с алгоритмом сегментации строк. Для распознавания личных данных используется система распознавания с учителем, реализующая метод «Ближайшего соседа». Обучение системы выполняется при запуске приложения, в результате чего формируется список классов из изображений эталонов букв латинского алфавита. Все данные, полученные в результате распознавания, передаются на форму системы управления, которая после проверки сохраняется в базе данных.

В системе управления используется две базы данных, одна для студентов, проживающих в общежитии, другая для студентов, утративших на постоянное место жительства. Обе базы созданы на платформе SQL Server 2005.

В данном проекте использована технология, основанная на стандарте COM – OLE (ObjectLinkingandEmbedding). Взаимодействие разработанной программы и объектной моделью текстового редактор Microsoft Word осуществляется с помощью сборки Microsoft.Office.Interop.Word.dll.

Программное приложение реализовано в среде программирования MS Visual Studio 2008 на языке C#, что обеспечивает более широкие возможности в работе с COM-объектами и базами данных.