

УДК 004.021

ИНТЕЛЛЕКТУАЛЬНОЕ ОБЪЕДИНЕНИЕ И КОМПОЗИЦИЯ ИЗОБРАЖЕНИЙ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Чекашев А.А.(Университет ИТМО)

Научный руководитель – магистр, инженер Ефимова В.А.
(Университет ИТМО)

В области графического дизайна часто необходимо подобрать изображение в качестве иллюстрации или обложки. Данный алгоритм позволит автоматически создавать изображения высокого разрешения, подходящие по контексту.

Введение. Важной задачей оформления книги является подбор изображений в качестве иллюстраций и обложки. Обычно в качестве иллюстраций и обложек используются готовые изображения или фото. Этот процесс может быть дорогим и трудоемким, особенно для начинающих авторов. Наличие у книги обложки делает ее более привлекательной для аудитории, а иллюстрации делают подачу материала более наглядной. Автоматизация этого процесса позволит авторам самостоятельно создавать подходящие иллюстрации, прикладывая лишь немного усилий. Генерация изображений по краткому описанию является популярной задачей в области машинного обучения. Однако множество современных решений базируются на нейронных сетях, которые не позволяют получить высокодетализированные изображения. К тому же, подобные алгоритмы обычно показывают хорошие результаты только для запросов с определенным контекстом, данные которого были использованы на стадии обучения.

Основная часть. Создавать изображение предлагается путем объединения нескольких изображений, подходящих под контекст. Для начала, из текста выделяются ключевые слова, определяющие тематику изображения. С использованием этих слов из базы выбираются изображения, которые впоследствии будут объединяться. Для каждого изображения методами машинного обучения выделяются объекты, которые на нем находятся. На основе этой информации каждое из изображений попадает либо в категорию изображений-фонов, либо в категорию изображений, из которых будут извлекаться объекты. Извлечение объектов происходит с помощью удаления фона. Полученные объекты группируются и располагаются на изображениях-фонах с учетом карт глубины. В конце процедуры на изображение наносятся визуальные эффекты, а также интеллектуально выбирается наилучший кадр необходимого формата.

Выводы. Данный подход позволит получать изображения высокого разрешения при высоком разрешении изображений из базы. При высокой скорости работы, пользователю будут доступны несколько результатов, из которых он сможет выбрать тот, который его устраивает. Проверка корректности алгоритма может быть произведена путем опроса пользователей, в котором опрашиваемый должен выбрать наиболее подходящее изображение по заданному тексту. Среди предложенных изображений должны быть как изображения дизайнеров, так и изображения, полученные с помощью алгоритма. Если в результате опроса некоторая часть пользователей будет отдавать предпочтение сгенерированному изображением, то можно считать, что проверка прошла успешно.

Чекашев А.А. (автор)

Подпись

Ефимова В.А. (научный руководитель)

Подпись