

ВСТАВКА ОБЪЕКТА В ИЗОБРАЖЕНИЕ С УЧЕТОМ РАЗНИЦЫ В ОСВЕЩЕНИИ

Овечкин В.Е. (Университет ИТМО)

Научный консультант – Ефимова В.А. (Университет ИТМО)

Научный руководитель – к.ф.-м.н., Фильченков А.А. (Университет ИТМО)

В данной работе рассматривается метод вставки объекта в изображение, использующий векторное представление освещения для выравнивания освещения объекта и фона. Таким образом можно добиться более реалистичного вписывания объекта в изображение, по сравнению с обычной вставкой.

Введение. Задача реалистичной вставки объекта в изображение стоит достаточно давно, тем не менее тема не теряет актуальности. В последнее время ежегодно публикуется по несколько работ по этой теме, но результаты все еще далеки от идеальных. В первых исследованиях работа велась в основном с низкоуровневыми признаками изображений, например с цветом пикселей. Современные методы используют такие высокоуровневые факторы, как положение и интенсивность источников света на изображении, время суток, геометрия сцены. Совокупность всех таких признаков и является освещением изображения. Если найти закономерности, по которым изменение параметров освещения влияет на цвет пикселей изображения, то можно будет научиться восстанавливать параметры освещения по изображению. В этом случае, узнав освещение фона, можно будет изменить освещение объекта таким образом, чтобы реалистично вписать его в фон.

Основная часть. Целью данной работы является реалистичная вставка объекта в изображение с учетом разницы в их освещении. Для решения этой задачи требуется узнать освещение как объекта, так и фона для того, чтобы впоследствии их выровнять. Освещение – сложная функция различных факторов изображения. В этой работе предлагается новый подход к оценке освещения изображения – представлять освещение в виде вектора (эмбединга). По аналогии с эмбедингами для текстов, эмбединг освещения отображает изображение в многомерное векторное пространство, где каждой точке соответствует какая-то функция освещения в реальном мире.

Для построения эмбединга к существующей нейронной сети ResNet101 были добавлены несколько независимых полносвязных слоев, каждый из которых аппроксимировал какой-то параметр изображения: наличие источников света, яркость, контрастность, насыщенность. Для обучения получившейся сети был использован набор данных The Laval Indoor HDR Dataset. Такие изображения хранят информацию о пикселе в 32-битном формате для каждого цветового канала, что позволяет добиться более точных значений освещенности изображения. Для каждого изображения подсчитывались целевые значения, после чего пары из изображений и меток передавались сети для обучения. Выход слоя обученной сети, предшествующего полносвязным слоям, и будет векторным представлением освещения.

Имея эмбединг освещения, можно отобразить объект и фон в две точки в векторном пространстве освещения. После чего нужно найти вектор разности этих точек. Он будет представлять собой разность в освещении объекта и фона. Теперь по этому вектору нужно научиться генерировать фильтр для объекта, который будет выравнивать его освещение по сравнению с фоном. Для решения этой задачи потребуется генеративная сеть, которой на вход будет подаваться эмбединг освещения. Она будет генерировать функцию фильтра освещенности для объекта, после чего к объекту будет применяться этот фильтр и он будет вставляться в фон. В качестве функции ошибки используется результат еще одной

нейронной сети RealismCNN, которая предсказывает реалистичность композиции двух изображений.

Обученную на множестве пар из фона и объекта сеть можно использовать для коррекции освещения при вставке нового объекта в изображение.

Выводы. Данная технология может быть использована широким кругом людей вместо редакторов изображений, так как не требует никаких дополнительных навыков. Также ее можно использовать для вставки рекламы в готовые видеоролики без дополнительного монтажа.

Овечкин В.Е. (автор)

Подпись

Ефимова В.А. (научный консультант)

Подпись

Фильченков А.А. (научный руководитель)

Подпись