

Приложение F.

Инструкция асессора

для дорожки поиска нечетких

дубликатов в коллекции

изображений

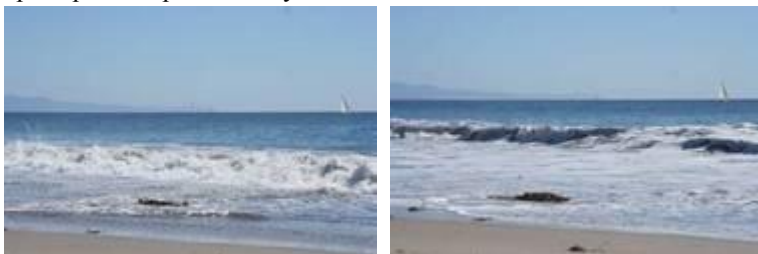
F.1 Постановка задачи

Ваша задача - оценить качество поиска дубликатов в коллекции изображений.

Каждое задание состоит из набора картинок, среди которых Вам необходимо выделить группы дублей (до 25 групп в одном наборе). Дубликатами считаются фотографии одной и той же сцены или объекта, сделанные в разных условиях, или разного качества. В частности, дубликатами являются фотографии, снятые в разном масштабе или с разных точек, с различиями в фокусном расстоянии, освещении, с незначительными изменениями фона (движение волны в море или листьев на дереве).

F.2 Примеры

Примеры изображений-дубликатов:





Примеры визуально и/или семантически похожих изображений, не являющихся при этом дублями:



Выделение групп дублей осуществляется путем присвоения картинкам номеров групп: картинкам-дубликатам необходимо присвоить один и тот же номер группы. Ниже приведены основные правила работы с инструментах по оценке.

- Цель оценки: выделить до 25 групп похожих картинок (дублей) в заданном наборе.
- Для выделения группы дублей необходимо выделить с помощью мышки/клавиатуры картинки, относящиеся к одной группе, и нажать на кнопку с номером группы.

- Группы картинок можно пополнять путем выделения новых картинок, принадлежащих уже существующей группе, и нажатия на номер группы, которую хотим пополнить.
- Созданные группы можно редактировать и просматривать, выбрав из списка созданных групп, расположенном на правой панели.
- Не обязательно присвоить номер группы каждому из изображений набора.
- Не надо создавать группы из одного изображения. Можно просто оставить их неоцененными.
- Начинайте выделять с крупных групп. Мы хотим оценить как можно больше изображений.
- Оценивать удобней, если окно инструмента достаточно большое — так видно больше изображений. Рекомендуем развернуть инструмент на весь экран.
- Любые два изображения из выделенной группы должны быть дублями.

Во многих случаях вы будете видеть ряд изображений, где соседние изображения очень похожи, но постепенно накапливаются значительные отличия. Например, сначала несколько изображений машины на фоне моря, потом кусочка машины на фоне моря, а потом просто моря. Будет неправильно отнести их в одну группу и может быть несколько вариантов как разбить их на группы (возможно оставив часть неоцененными). Решение стоит принимать исходя из того как вам кажется правильно было бы их сгруппировать.

- Часть изображений имеет плохое качество. Изображение с плохим и хорошим качеством могут быть дублями, если они попадают под данные выше определения. Качество изображения не является их отличающим признаком.