

Комплексное руководство по аннотированию POV-видео ручной работы

Оглавление

1. Введение и основные принципы
 2. Структура описания действий
 3. Система предлогов и семантических связей
 4. Временная сегментация
 5. Контролируемый словарь
 6. Протокол предварительной разметки экспертом
 7. Правила качественной аннотации
 8. Практические примеры
 9. Чек-листы для работы
 10. Частые ошибки и как их избежать
-

1. Введение и основные принципы {#введение}

Цель данного руководства

Обеспечить создание аннотаций, которые:

- Могут быть автоматически обработаны ML-моделями
- Содержат достаточную семантическую информацию для обучения
- Следуют единому стандарту для обеспечения консистентности
- Максимально используют возможности однослойной системы аннотирования

Ключевые принципы

1. **Консистентность важнее объёма** - лучше 1000 идеально стандартизированных аннотаций, чем 5000 с вариациями
 2. **Плотность важнее покрытия** - лучше тщательно аннотировать 30 минут, чем поверхностно 2 часа
 3. **Структура в рамках ограничений** - используем шаблоны для кодирования многослойной информации в однослойной системе
-

2. Структура описания действий {#структура-описания}

Базовый шаблон

[ГЛАГОЛ] + [ОБЪЕКТ/ОБЛАСТЬ] + [ИНСТРУМЕНТ] + [ДЕТАЛИ]

Уровни сложности

Уровень 1 - Минимальный (редко используется):

НАНЕСТИ лак на_ноготь_большого_пальца

Уровень 2 - С инструментом (80% всех аннотаций):

НАНЕСТИ лак на_ноготь_большого_пальца кисточкой

Уровень 3 - С деталями (15% аннотаций):

НАНЕСТИ лак на_ноготь_большого_пальца кисточкой (тонкий_слой)

Уровень 4 - С указанием руки (5% аннотаций, только когда критично):

НАНЕСТИ лак на_левый_большой_палец кисточкой_в_правой_руке (тонкий_слой)

Правило 80/20

- 80% аннотаций должны быть уровня 2
- 15% аннотаций должны быть уровня 3
- 5% аннотаций должны быть уровня 4

3. Система предлогов и семантических связей {#система-предлогов}

Зачем использовать подчёркивания

Подчёркивание создаёт **атомарные семантические единицы**:

- Машинная обработка воспринимает их как единый токен
- Предотвращает неоднозначность при парсинге
- Связывает предлог с объектом в единое понятие

Детальные правила предлогов

С_ (инструмент/средство)

Значение: указывает инструмент или средство выполнения действия

Шаблон: ([ГЛАГОЛ] [объект] с_[инструмент])

Примеры:

ПОДПИЛИТЬ ногти с_пилочкой
ОЧИСТИТЬ кутикулу с_деревянной_палочкой
УДАЛИТЬ лак с_ватным_диском
ПРИДАТЬ_ФОРМУ ногтям с_кусачками
ПОЛИРОВАТЬ поверхность с_четырёхсторонним_бафом

НА_ (цель/направление)

Значение: указывает, куда что-то наносится, перемещается или направляется

Шаблон: ([ГЛАГОЛ] [объект/материал] на_[цель])

Примеры:

НАНЕСТИ базовое_покрытие на_все_ногти
ПЕРЕМЕСТИТЬ лампу на_левую_сторону
ПЕРЕНЕСТИ лак на_палитру
ДОБАВИТЬ блёстки на_безымянный_палец
ПРИКРЕПИТЬ стразу на_кончик_ногтя

С (источник/происхождение)

Значение: указывает, откуда что-то удаляется или берётся

Шаблон: ([ГЛАГОЛ] [объект] с_[источник])

Примеры:

УДАЛИТЬ лак с_большого_пальца
ВЗЯТЬ пилочку с_стола
СТЕРЕТЬ пыль с_поверхности_ногтя
ВЫБРАТЬ цвет с_палитры

ДЛЯ_ (цель/длительность/получатель)

Значение: указывает зачем, как долго или для кого что-то делается

Шаблон: [ГЛАГОЛ] [объект] для [цель/время/персона]

Примеры:

ЖДАТЬ для_высыхания_лака
ПОДГОТОВИТЬ ногти для_нанесения_геля
ПРОВЕРИТЬ ногти для_выявления_пузырьков
ЗАМОЧИТЬ пальцы для_5_минут
ВЫБРАТЬ цвет для_клиента

В_ (контейнер/состояние/положение)

Значение: указывает на содержание, состояние или положение

Примеры:

ОКУНУТЬ кисточку в_бутылку_с_лаком
ПОМЕСТИТЬ инструменты в_стерилизатор
ЗАМОЧИТЬ ногти в_растворе
СУШИТЬ ногти в_LED_лампе

Сложные комбинации

Можно комбинировать несколько предлогов для богатой семантики:

НАНЕСТИ гель_лак на_большой_палец с_тонкой_кисточкой для_базового_слоя
УДАЛИТЬ излишки_кутикулы с_ногтевого_ложа с_помощью_кусачек
ПЕРЕНЕСТИ дизайн с_трафарета на_безымянный_палец с_помощью_штампа

4. Временная сегментация {#временная-сегментация}

Метод критических точек контакта

Основан на системе PRE-CONTACT-POST из Ego4D:

Правила временных границ:

- **Начало** = момент движения руки к инструменту/объекту (не контакт!)
- **Конец** = момент отпускания объекта ИЛИ начало движения к следующему инструменту
- **Минимальная длительность** = 2 секунды (более короткие действия объединяем)
- **Максимальная длительность** = 60 секунд (более длинные разбиваем на естественных паузах)

Пример сегментации

00:15-00:22 ПОДГОТОВИТЬ ноготь с_пилочкой (удаление_старого_лака)
00:22-00:28 ВЗЯТЬ бутылку_ацетона
00:28-00:35 НАНЕСТИ ацетон на_ватный_диск
00:35-00:48 УДАЛИТЬ лак с_всех_ногтей с_дискон
00:48-00:52 ВЫБРОСИТЬ ватный_диск

Иерархия в плоской структуре

Используйте префиксы для кодирования уровней:

- **ЭТАП:** для высокоуровневых фаз (3-5 минут)
- **ДЕЙСТВИЕ:** для функциональных единиц (10-60 секунд)
- **МИКРО:** для критически важных деталей (2-10 секунд)

Пример:

00:00-03:45 ЭТАП: подготовка_ногтей
00:00-00:35 ДЕЙСТВИЕ: удаление_старого_лака (все_пальцы)
00:35-00:48 ДЕЙСТВИЕ: очистка_поверхности с_ацетоном
00:48-00:52 МИКРО: проверка_ногтей на_остатки

5. Контролируемый словарь {#контролируемый-словарь}

Основные глаголы для ручной работы (ограничить 30-50)

Подготовка:

- ПОДГОТОВИТЬ, ОЧИСТИТЬ, ДЕЗИНФИЦИРОВАТЬ, ОРГАНИЗОВАТЬ, ВЫБРАТЬ

Манипуляция:

- ВЗЯТЬ, ПОЛОЖИТЬ, ДЕРЖАТЬ, РАСПОЛОЖИТЬ, ПОВЕРНУТЬ

Нанесение:

- НАНЕСТИ, РАСПРЕДЕЛИТЬ, ПОКРЫТЬ, НАКРАСИТЬ, РАСПЫЛИТЬ

Модификация:

- ПОДРЕЗАТЬ, ПОДПИЛИТЬ, ПРИДАТЬ_ФОРМУ, ОБРЕЗАТЬ, ПОЛИРОВАТЬ

Удаление:

- УДАЛИТЬ, СТЕРЕТЬ, ИЗВЛЕЧЬ, СНЯТЬ, СОСКОБЛИТЬ

Проверка:

- ПРОВЕРИТЬ, ОСМОТРЕТЬ, ИЗУЧИТЬ, СРАВНИТЬ

Ожидание:

- ЖДАТЬ, СУШИТЬ, ЗАКРЕПИТЬ, ПОЛИМЕРИЗОВАТЬ

Таксономия объектов/областей

Конкретные части:

- большой_палец, указательный_палец, кутикула, ногтевое_ложе

Группы:

- все_ногти, левая_рука, правая_рука

Материалы:

- лак, ацетон, масло, крем, гель

Инструменты:

- пилочка, баф, пушер, кусачки, лампа

Дескрипторы состояния (в скобках)

- (первый_слой), (второй_слой), (финальный_слой)
- (левая_завершена), (правая_завершена)
- (крупное_зерно), (мелкое_зерно)
- (толстое_нанесение), (тонкое_нанесение)

6. Протокол предварительной разметки экспертом {#протокол-эксперта}

Шаблон для голосовых заметок эксперта

Эксперт записывает описание каждые 5 минут:

"Минуты 0-5: Я подготавливаю натуральные ногти - удаляю старый лак, отодвигаю кутикулу и придаю форму по желанию клиента - круглые кончики"

"Минуты 5-10: Базовая подготовка - обезжириваю ногти, наношу праймер, первый слой базы с акцентом на запечатывание кончиков"

Преобразование в маркеры ЭТАП

00:00-05:00 ЭТАП: подготовка_натуральных_ногтей [удаление_лака, работа_с_кутикулой, придание_формы]

05:00-10:00 ЭТАП: базовая_подготовка [обезжиривание, праймер, база]

7. Правила качественной аннотации {#правила-качества}

Метод трёх проходов

Проход 1 - Структура (смотрим на скорости 2х):

- Отмечаем границы ЭТАПОВ каждые 3-5 минут
- Замечаем смены инструментов
- Идентифицируем основные переходы

Проход 2 - Действия (смотрим на нормальной скорости):

- Заполняем аннотации уровня ДЕЙСТВИЕ
- Фокусируемся на взятии/откладывании инструментов
- Отмечаем все изменения состояния объектов

Проход 3 - Уточнение (прокручиваем конкретные секции):

- Добавляем МИКРО аннотации для сложных техник
- Проверяем временные границы
- Обеспечиваем консистентность словаря

Обработка неактивности и переходов

Продуктивное ожидание (оставляем как активное):

02:15-02:45 ЖДАТЬ для_высыхания_лака (первый_слой)

Истинная неактивность (помечаем как неактивное):

02:45-03:10 [НЕАКТИВНО] разговор_с_клиентом

Маркеры перехода:

05:30-05:35 ПЕРЕХОД переключение_на_правую_руку
10:15-10:20 ПЕРЕХОД подготовка_следующего_цвета

Стратегии максимизации ценности

Плотность превыше покрытия:

- Минимум 1-2 аннотации в минуту

- Включайте неудачные попытки: **НАНЕСТИ лак на_большой_палец (неудачное_покрытие)**

Фиксируйте исправления и вариации:

03:15-03:25 УДАЛИТЬ лак с_большого_пальца (исправление_ошибки)
03:25-03:35 ПОВТОРНО_НАНЕСТИ лак на_большой_палец (вторая_попытка)

Связывайте родственные действия через ID:

01:15-01:25 ПОДГОТОВИТЬ большой_палец (подготовка_1)
01:25-01:35 ПРИДАТЬ_ФОРМУ большому_пальцу (после_подготовка_1)
05:15-05:25 ПОКРЫТЬ большой_палец (после_подготовка_1)

8. Практические примеры {#примеры}

Пример 1: Базовый маникюр (2 минуты)

00:00-00:08 ВЗЯТЬ пилочку (среднее_зерно)
00:08-00:22 ПОДПИЛИТЬ большой_палец (придание_круглой_формы)
00:22-00:35 ПОДПИЛИТЬ указательный_палец (придание_круглой_формы)
00:35-00:48 ПОДПИЛИТЬ средний_палец (придание_круглой_формы)
00:48-00:55 ПРОВЕРИТЬ все_ногти (консистентность_формы)
00:55-01:02 ПОЛОЖИТЬ пилочку
01:02-01:08 ВЗЯТЬ баф (гладкая_поверхность)
01:08-01:20 ПОЛИРОВАТЬ большой_палец (удаление_неровностей)
01:20-01:32 ПОЛИРОВАТЬ указательный_палец (удаление_неровностей)
01:32-01:45 ПОЛИРОВАТЬ средний_палец (удаление_неровностей)
01:45-01:52 ОСМОТРЕТЬ левую_руку (качество_поверхности)
01:52-02:00 ПЕРЕХОД переключение_на_правую_руку

Пример 2: Сложная техника с указанием рук

00:00-00:10 СТАБИЛИЗИРОВАТЬ левый_указательный левой_рукой
00:00-00:10 НАНЕСТИ полосу_для_дизайна на_левый_указательный правой_рукой
00:10-00:15 ПРИЖАТЬ полосу с_деревянной_палочкой для_адгезии
00:15-00:25 НАНЕСТИ лак поверх_полоски с_кисточкой

Пример 3: Работа с гель-лаком

00:00-00:15 ЭТАП: подготовка_к_гель_лаку
00:00-00:05 ОБЕЗЖИРИТЬ все_ногти с_дегидратором
00:05-00:10 НАНЕСТИ праймер на_все_ногти

00:10-00:15 ЖДАТЬ для_высыхания_праймера (30_секунд)
00:15-00:45 ДЕЙСТВИЕ: нанесение_базы
00:15-00:20 ВЗЯТЬ базовый_гель
00:20-00:35 НАНЕСТИ базу на_все_ногти с_кисточкой (тонкий_слой)
00:35-00:40 ЗАПЕЧАТАТЬ торцы с_кисточкой
00:40-00:45 ПОМЕСТИТЬ руку в_LED_лампу
00:45-01:15 ПОЛИМЕРИЗОВАТЬ базу в_лампе (30_секунд)

9. Чек-листы для работы {#чек-листы}

Перед началом работы

- ☐ Создать список глаголов (максимум 50)
- ☐ Создать список объектов/инструментов для вашей области
- ☐ Получить высокоуровневое описание от эксперта
- ☐ Определить правила минимальной/максимальной длительности
- ☐ Создать шпаргалку с шаблонами

Во время аннотирования

- ☐ Начать с маркеров ЭТАП из описания эксперта
- ☐ Аннотировать смены инструментов как опорные точки
- ☐ Использовать консистентный паттерн глагол_объект_инструмент
- ☐ Различать неактивность и продуктивное ожидание
- ☐ Отмечать любые неоднозначные случаи для ревью

Контроль качества

- ☐ Нет перекрывающихся временных меток
- ☐ Все описания следуют шаблону
- ☐ Минимум 1 аннотация в минуту
- ☐ Все глаголы из контролируемого списка
- ☐ Смены инструментов явно отмечены

Финальная проверка

- ☐ Временная логика соблюдена (действия не начинаются раньше предыдущих)
 - ☐ Пространственная консистентность (боксы не прыгают больше чем на 50 пикселей)
 - ☐ Валидность меток (термины из контролируемого словаря)
 - ☐ Плотность аннотаций достаточная (1-2 в минуту)
-

10. Частые ошибки и как их избежать {#ошибки}

Ошибка 1: Неконсистентная маркировка

❌ **Неправильно:** использовать "кот", "кошка", "котёнок" для одного объекта ✅ **Правильно:** выбрать один термин и использовать его везде

Ошибка 2: Свободный текст вместо структуры

❌ **Неправильно:** "наносит лак на ноготь используя кисточку" ✅ **Правильно:** НАНЕСТИ лак на _ноготь с _кисточкой

Ошибка 3: Пропуск инструментов

❌ **Неправильно:** ПОДПИЛИТЬ ногти ✅ **Правильно:** ПОДПИЛИТЬ ногти с _пилочкой

Ошибка 4: Неточные временные границы

❌ **Неправильно:** начинать с момента контакта ✅ **Правильно:** начинать с момента движения к объекту

Ошибка 5: Излишняя детализация

❌ **Неправильно:** ВЗЯТЬ пилочку (синяя _ручка) правой _рукой под _углом _45 _градусов ✅ **Правильно:** ВЗЯТЬ пилочку

Ошибка 6: Игнорирование неудач

❌ **Неправильно:** пропускать неудачные попытки ✅ **Правильно:** НАНЕСТИ лак на _ноготь (неудачная _попытка)

Ошибка 7: Использование местоимений

❌ **Неправильно:** "почистить его" ✅ **Правильно:** ОЧИСТИТЬ большой _палец

Ошибка 8: Несвязанные действия

❌ **Неправильно:** не показывать связь между подготовкой и результатом ✅ **Правильно:** использовать ID для связи: ПОДГОТОВИТЬ ноготь (подг_1), затем ПОКРЫТЬ ноготь (после_подг_1)

Дерево принятия решений для выбора уровня детализации

Действие требует указания инструмента?

└─ ДА → Добавить с_[инструмент]

| └─ Техника важна для результата?

| └─ ДА → Добавить (детали)

| └─ Руки выполняют разные функции?

| └─ ДА → Указать какая рука что делает

| | └─ НЕТ → Оставить без указания руки
| └─ НЕТ → Оставить без деталей
└─ НЕТ → Базовая аннотация

Пример парсинга для анализа

Ваши структурированные описания можно программно обрабатывать:

```
python

# Пример логики парсинга
description = "ПОДПИЛИТЬ большой_палец с_пилочкой (придание_формы)"
verb = description.split()[0] # "ПОДПИЛИТЬ"
object = description.split()[1].split("_")[0] # "большой_палец"
tool = description.split("_")[1].split("(")[0] if "_" in description else None
detail = description.split("(")[1].strip("(") if "(" in description else None
```

Это позволяет генерировать:

- Анализ частотности глаголов
- Паттерны использования инструментов
- Последовательности взаимодействия с объектами
- Статистику времени на действие

Ключевой принцип успеха

Несмотря на наличие только однослойной аннотации, данный подход обеспечивает:

1. **Композиционную структуру** через стандартизированные шаблоны (обеспечивает zero-shot обучение)
2. **Иерархическую информацию** через префиксы (ЭТАП/ДЕЙСТВИЕ/МИКРО)
3. **Временную точность** через границы на основе контактов
4. **Семантическое богатство** через контролируемый словарь + детали
5. **Гарантию качества** через парсируемые паттерны

Исследования показывают, что аннотации с консистентной структурой и контролируемым словарём улучшают производительность моделей на 15-25% по сравнению со свободным текстом, даже при простых инструментах.

Заключение

Критический фактор успеха - **консистентность**. Лучше иметь 1000 аннотаций, идеально следующих этим паттернам, чем 5000 с вариациями. Тщательно обучите аннотаторов этим шаблонам и регулярно проверяйте соответствие паттернам.

Данное руководство обеспечивает баланс между:

- Достаточной детализацией для воспроизведения техники
- Простотой для лёгкого извлечения паттернов
- Гибкостью для адаптации под разные типы ручной работы

Следуя этим принципам, вы создадите аннотации, которые будут ценны как для текущих задач машинного обучения, так и для будущих исследований в области компьютерного зрения и робототехники.