

如何制作大规样本的 `ImgMat` 数据库

在 `zAI_Common.pas` 中，`TAI_ImageMatrix` 代表大规模数据集

通过申明，你会发现 `TAI_ImageMatrix` 实际上是一个泛型链表

`TAI_ImageMatrix` 动态数组链里面的元素则是 `TAI_ImageList`，`TAI_ImageList` 这是 `.imgDataset` 的驱动类

换句话说，`TAI_ImageMatrix` 是一个 `.imgDataset` 数据集仓库，在 `zAI` 中，这是以大规模存储为目标而编写的。

动手提炼 `.imgDataset`，通过 `demo`，`AISetFormatExtract`，我们可以了解到手动提炼过程的来龙去脉

当我们提炼完成 `.imgDataset` 后，将 `TAI_ImageList` 以元素方式添加到 `TAI_ImageMatrix` 即完成大规模数据库。

`TAI_ImageList.create`

`TAI_ImageList` 的提炼过程程序

完成 `TAI_ImageList` 后 `TAI_ImageMatrix.add(TAI_ImageList)`

`TAI_ImageMatrix.LargeScale_SaveToFile(...)`

通过阅读 `Demo` 和分析申明头，只要弄清楚数据结构，制作会很容易。更多的难点是结合图像方便的概念来编程。

针对特大 `TAI_ImageList` 模型的说明：在使用 `LargeScale_SaveToFile` 方法时，`TAI_ImageList` 会使用 4 倍存储 `size` 的临时内存开销，简单来说，`TAI_ImageList` 如果使用了 2G 内存，在使用 `LargeScale_SaveToFile` 时，临时内存开销需要 8G，用完即刻被释放。`TAI_ImageMatrix` 并不会对单一而又巨大的 `TAI_ImageList` 进行 `LargeScale` 处理，对于大批小尺寸的 `TAI_ImageList` 才有 `LargeScale` 效果。

2019-7

By.qq600585