Recently, sparse representation based generative tracking methods have been developed for object tracking [1, 2, 3, 4, 5, 6, 7, 9, 10, 11]. These trackers can be categorized based on the representation schemes into global, local, and joint sparse appearance models as shown in Figure 1. There exist three models [1, 2, 3, 4, 5, 6, 7, 9, 10, 11]. Using our model, all particles \( X \) and their local patches are represented with joint sparsity by the corresponding local patches inside different dictionary templates. (c) Joint sparse appearance model [1, 10, 11]. These methods are grouped based on their sparse appearance models.

The comparison results on benchmark [8] are shown in Figure 3. The results show that our SST tracker achieves favorable performance than other related sparse trackers [2, 6, 7, 9, 10]. Compared with other state-of-the-art methods, our SST achieves the second best overall performance.

In this paper, we use the convex \( \ell_{p,q} \) mixed norm, especially, \( \ell_{2,1} \) to model the structure information of \( Z^k \) and \( Z_q \) and obtain the structural sparse appearance model for object tracking as

\[
\min_{Z} \frac{1}{2} \sum_{k=1}^{K} \|X^k - D^kZ^k\|_F^2 + \lambda \|Z\|_{2,1},
\]

where \( Z = [Z^1, Z^2, \ldots, Z^K] \in \mathbb{R}^{m \times nK} \), \( \| \cdot \|_F \) denotes the Frobenius norm, and \( \lambda \) is a tradeoff parameter between reliable reconstruction and joint sparsity regularization. The definition of the \( \ell_{p,q} \) mixed norm is \( \|Z\|_{p,q} = \left( \sum_{i} \left( \sum_{j} |Z_{ij}|^q \right)^{\frac{p}{q}} \right)^{\frac{1}{p}} \), and \( Z_{ij} \) denotes the entry at the \( i \)-th row and \( j \)-th column of \( Z \). Figure 2 illustrates the structure of the learned matrix \( Z \).

The comparison results on benchmark [8] are shown in Figure 3. The results show that our SST tracker achieves favorable performance than other related sparse trackers [2, 6, 7, 9, 10]. Compared with other state-of-the-art methods, our SST achieves the second best overall performance.


