Seeing the Arrow of Time

Pulkit Bansal
Manish Nalla
Rajat Jain
Ramamurthi Kumar

Aim:

- To explore whether we can observe Time’s Arrow in a temporal sequence?
- To examine what motions help discriminate forwards from backwards time.

Introduction:

- Videos of certain motions can be time-reversed without looking wrong.
- While there are some motions that look absurd on time-reversal.
- Low level visual information,that is, the underlying physics, is used to see Time’s Arrow, not object level cues.

Three Methods:

- Statistical Flow Method,
- Motion Causation Method,
- Auto Regression Method.

Data Set:

- Dataset consists of 180 video clips from YouTube, which was obtained manually using more than 50 keywords, like dance, etc.

- The dataset contains a few clips from “backwards” videos on YouTube.

- Discarded clips from professional studios, and that had poor focus, lighting, excessive camera motion, computer graphics, etc.

- Another dataset, “Tennis Ball Dataset”, comprises of 13 HD videos.

Statistical Flow Method

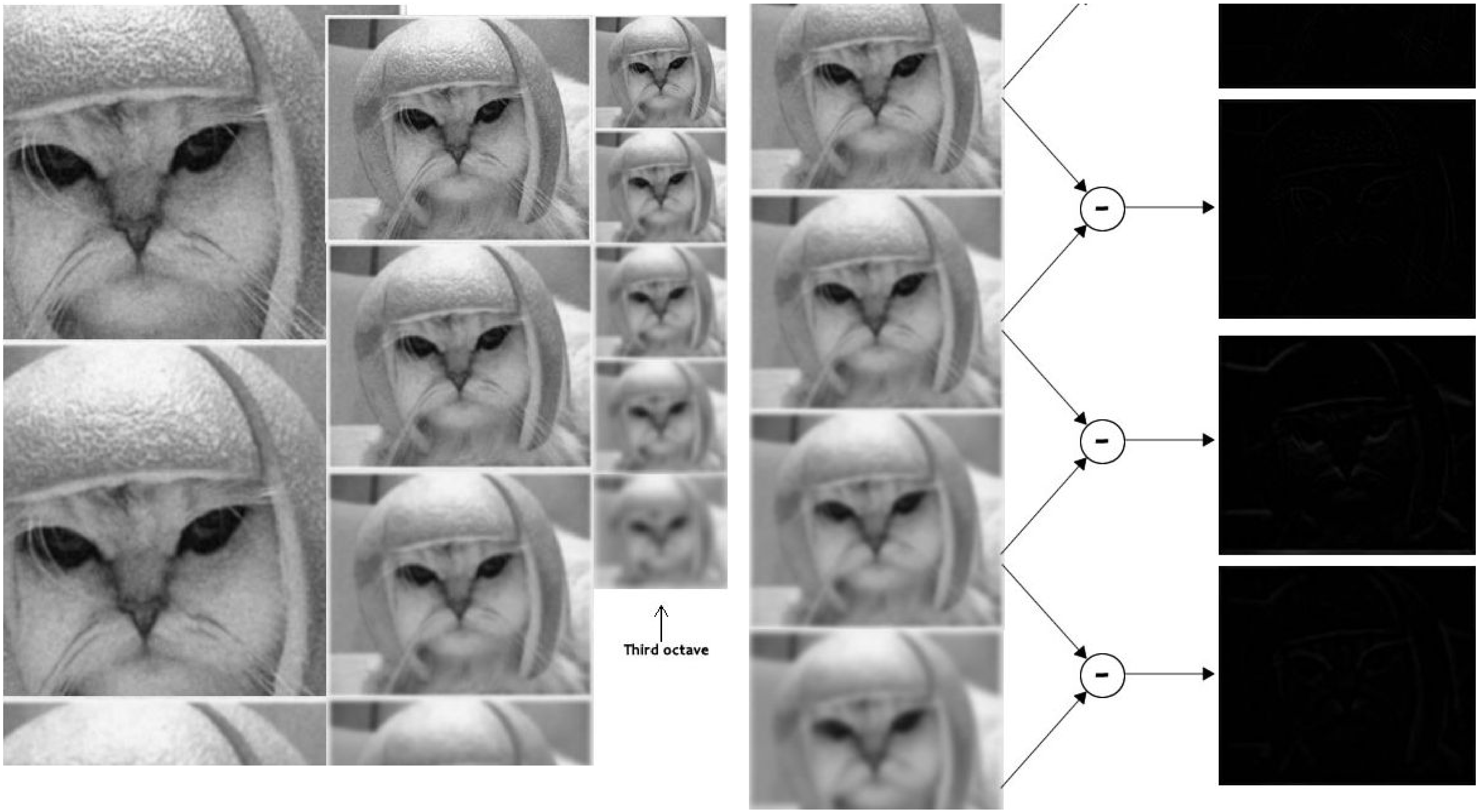
- Instead of using semantic information, patterns of motions are observed.
- Few reasons can differentiate forward and backward videos.
- Based on FLOW WORDS.
- Finding motions which exhibit temporal asymmetries.

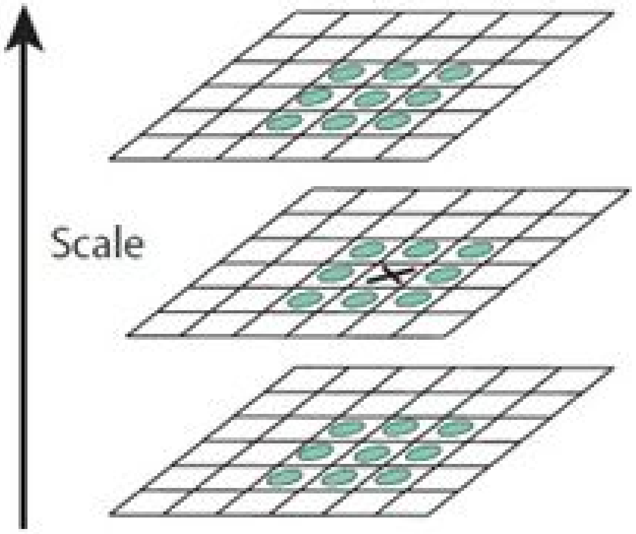
The Process(SFM):

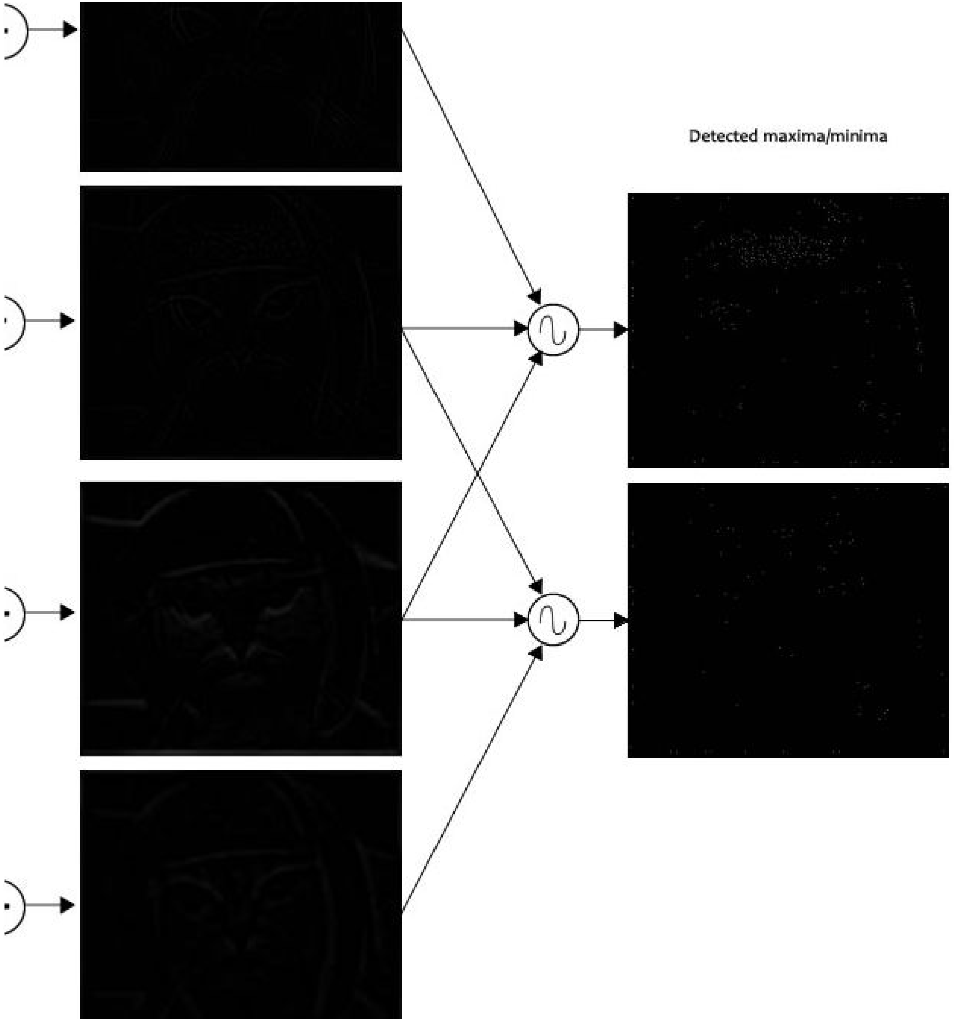
- Downsampling of images
- Registering images
- Optical Flow Outputs generated.
- SIFT-like descriptors used.

How the SIFT for Optical Flow works?

- Find Scale Space
- DoG approximation
- Finding KeyPoints
- Comparing Grids and measuring movement


Third octave


Scale


Detected maxima/minima

Algorithm(SFM):

- K-means Clustering to obtain flow words.
- 4 different types of flips performed for each video.
- SVM trained on the data set with labels.
- Different C values tested and the best one used.

Conclusion and Further Analysis:

- During the testing phase, for every video we find out the top 10 most informative Flow Words out of the 4000 clustered.
- These words' clusters were then marked to be forward/backward.
- Dataset divided into 3 and accuracy of 90%,77% and 70% obtained.

Motion Causation Method

- Based on cause-effect model
- Example :

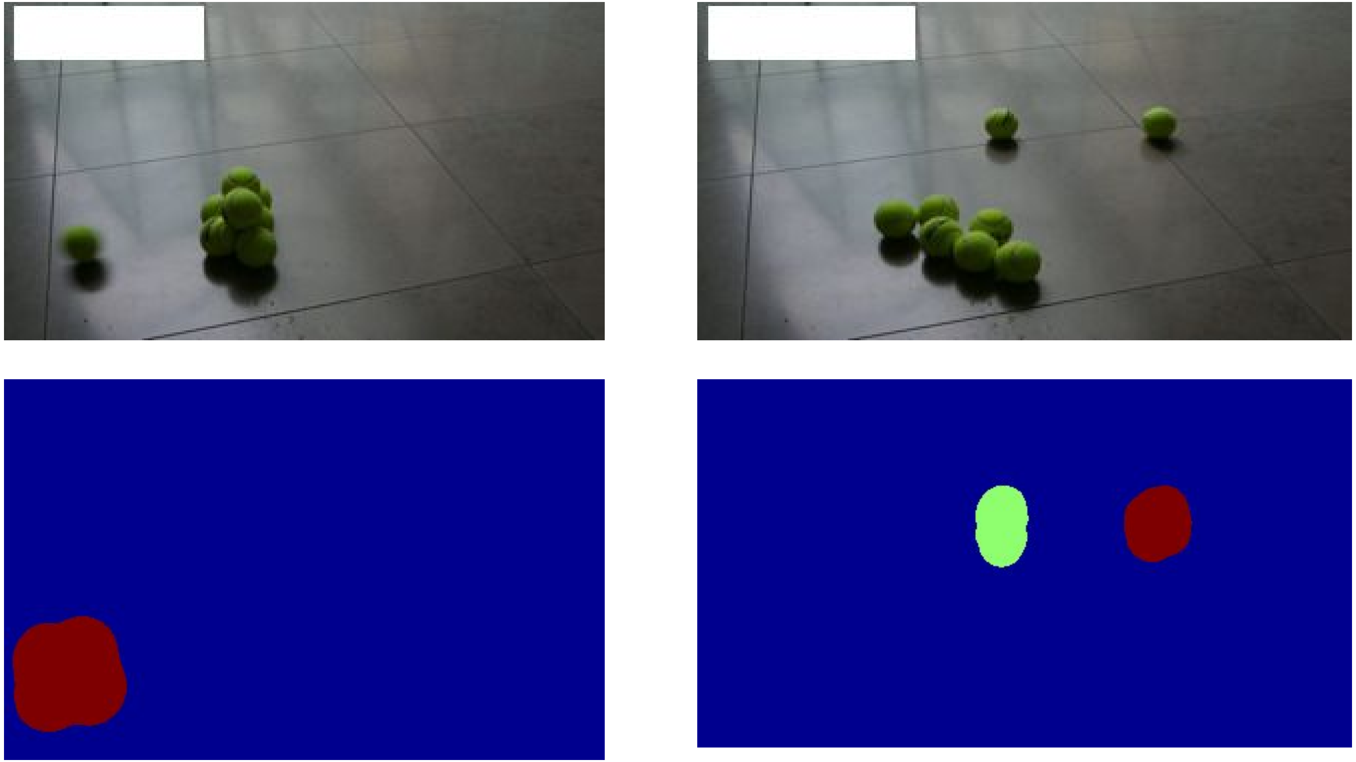
Frame 1

Frame 4

- Image warping :
    - Warp an image from a frame into another frame, to fill the regions.
    - The difference between this warped and original frame gives the moving areas.

- Noise cancellation
    - Using disk convolution
    - Using 3 Threshold radii pairs

ILLUSTRATION



- Count the violations
    - One or more regions at t intersect more than one region at t - 1, means its a violation.

- Train a SVM, for these violation counts as 6d features(for both the directions, three parameter settings)

- Obtain the results for the three video sets.

AUTO REGRESSION METHOD

*) measuring the direction of time as a special case of the problem of inferring causal direction in cause-effect models(usually based on time series model).

*) time series $(Xt)t \in Z$ an autoregressive model(AR model)

*) A time series $(X_t)_{t \in Z}$ is called **stationary** - distribution of a random vector $(X_{t1+h}, \ldots, X_{tn+h})$ does not change for any value of h. It is called **weakly stationary** if the mean is constant: $EX_t = \mu$ and the covariance function only depends on the time gap:

$$\mathrm{cov}(X_t, X_{t+h}) = \gamma h \ \forall t, h \in Z.$$

*) **DESIGN PRINCIPLE:** continuous random variables X and Y are related according to Y = $\alpha(Y_{t-1}) + \beta(Y_{t-2})+\varepsilon$, where ε is additive Noise. Regress both X on Y(forward direction) and Y on X(backward direction), and test for which direction the obtained noise is independent of the predictor variable: when the residual noise ε is independent of X, we determine X to be the cause and Y to be the effect. We will refer to this principle as causal direction through noise dependence.

*) x or y are velocities . used AR(2) means lag2

*) **Assumptions:**

Noise should be non gaussian

Linear Dynamics

*) the errors themselves follow a simple linear regression model that can be written as $\epsilon_t = b\epsilon_{t-1} + \omega$

*) Autocorrelation(t_1,t_2)= $E[(X_t\text{-mean}_t)(X_s\text{-mean}_s)]/(sd_t.sd_s)$ (x is velocity)(same for error)

*)for non-Gaussian additive noise and Linear dynamics AR model, the noise added at some point of time is independent of the past values of the time series, but not of the future values.

*)FLOW:
1) extract by KLT trackers running tracking in both forward and backward directions.
2)For each tracked point, velocities are extracted, and a 2D AR model is fitted.
3)We then test the independence between the noise and velocity

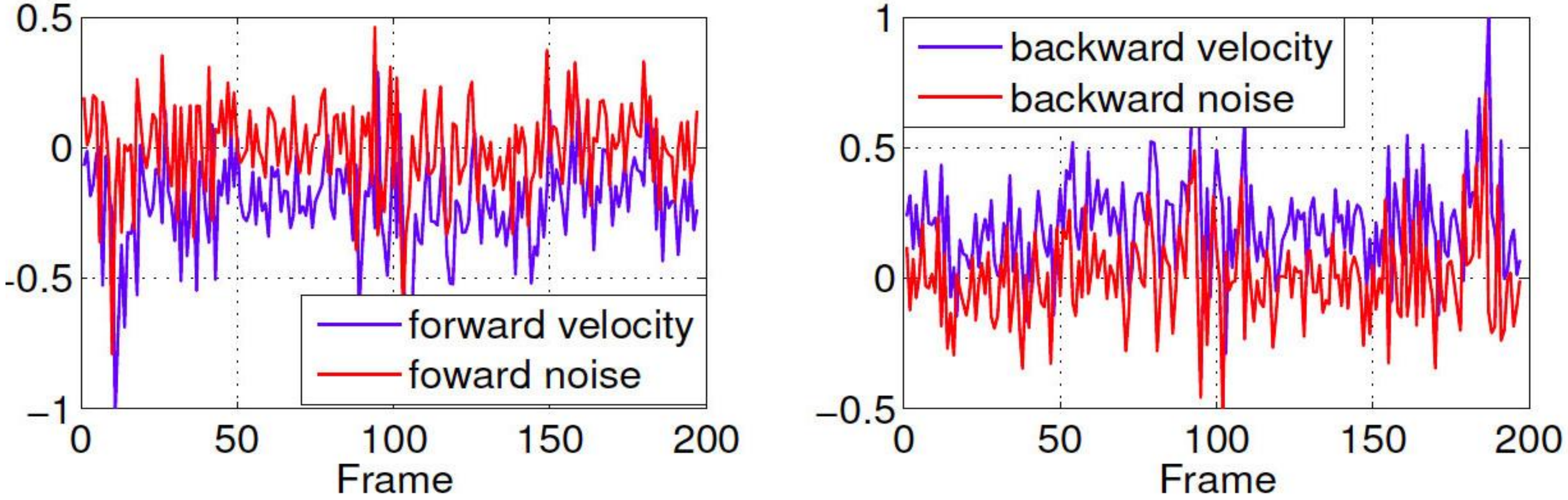
*)after independence task we get different p-values: pa and pb. If pa = 0.003 and pb = 0.43, for example, we reject the inde- pendence in the first case and we do not reject it in the second case.

RESULT

Whenever the noise is independent of the previous samples for one ordering of the observations, but dependent for the opposite ordering, we infer the former direction to be the true one.

ALGORITHM

```
input: a = (x1,...,xn), b = (xn,...,x1)
model_a =arfit(a)
res_a = model_a.residuals
model_b =arfit(b)
res_b = model_b.residuals
if res_a normally distributed then
        output =“I do not know (Gaussian process)”
        break
end if
if (res_a, a) independent then
        if (res_b, b) dependent then
                output =“(x1, . . . , xn) correct time direction”
        end if
else if (res_a, a) dependent then
        if (res_b, b) independent then
                output =“(xn, . . . , x1) correct time direction”
        else if (res_b, b) dependent then
                output =“I do not know (bad fit)”
end if
end if
```


0.5
0
-0.5
−1
0
50
100
150
200
Frame
forward velocity
foward noise
1
0.5
0
−0.5
0
50
100
150
200
Frame
backward velocity
backward noise

Conclusion:

- Results indicate that the statistics of natural videos are not symmetric under time reversal.

- Certain improvements can boost the performance further.