\title{Nanometer-scale bioimaging via tunable terahertz plasmons}

\author{Hasan Tahir Abbas}

\address{Department of Electrical and Computer Engineering,

Texas A$\&$M University, College Station, TX 77843 USA}

\address{Department of Electrical and Computer Engineering,

Texas A$\&$M University at Qatar, Doha, Qatar}

\author{Xiaodong Zeng}\email{zengxdgood@163.com}

\address{Institute for Quantum Science and Engineering (IQSE) and Department of Physics and Astronomy,

Texas A$\&$M University, College Station, Texas 77843-4242, USA}

\author{Jieyu You}

\address{Institute for Quantum Science and Engineering (IQSE) and Department of Physics and Astronomy,

Texas A$\&$M University, College Station, Texas 77843-4242, USA}

\author{Robert D. Nevels}

\address{Department of Electrical and Computer Engineering,

Texas A$\&$M University, College Station, TX 77843 USA}

\author{M. Suhail Zubairy}\email{zubairy@physics.tamu.edu}

\address{Institute for Quantum Science and Engineering (IQSE) and Department of Physics and Astronomy,

Texas A$\&$M University, College Station, Texas 77843-4242, USA}

\begin{abstract}

A nanometer-scale structured illumination microscopy (SIM) scheme via tunable plasmons is proposed. The sample is placed on a transistor based semiconductor heterostructure where terahertz plasmons generated by a current-driven instability, illuminate it. Full coverage of the spatial frequency regime is obtained by tuning the plasmons through gate voltage control. Hence, it is possible to reconstruct an image with resolution up to two orders of magnitude beyond the diffraction limit. Due to the linear nature of the technique, only a weak illumination signal is required, therefore minimizing the likelihood of sample damage due to radiation and having potential applications in bioimaging.

\end{abstract}

In conventional wide-field fluorescence microscopy, the resolution of the system is half the wavelength of light due to the Abbe diffraction limit. With an ever growing need to image minuscule objects especially in life sciences, several super-resolution microscopies have been realized. Structured illumination microscopy (SIM) \cite{Gustafsson2000,Gustafsson2005,Zeng2015,Heintzmann1999, Heintzmann2006} is a wide-field technique in which a fine illumination pattern such as a sinusoidal standing wave is used to generate \emph{Moir$\acute{e}$ fringes} in the observed image \cite{Heintzmann1999, Heintzmann2006}. The high spatial frequency content is mathematically reconstructed from a series of images acquired by shifting the pattern. Using a non-linear version of SIM, theoretically unlimited resolution can be achieved \cite{Gustafsson2005}. However, high levels of illumination intensity are required, which can subject the sample to significant damage due to thermal effects. Moreover, it has been found that resolution beyond the classical diffraction limit by a factor greater than $2$, can be realized when an object is illuminated by surface plasmons \cite{Wei2010,Zeng2014}.

Current-drive plasmon instabilities have mainly been studied in the context of ionized gases for a long time \cite{Mikh2003}. An analogous activity leads to generation of plasmons in solid-state devices that has resulted in many interesting applications in the far-infrared frequency region \cite{Kempa1991,Dyakonov1993, Dyakonov1996, Popov2005, Otsuji2006, Dyakonov2005,Stern1967a, Allen1977,Hofstetter2002, Dyer\_2016, Wu2015}. More importantly and interestingly, the spatial frequency response of the device can be tuned by varying the gate voltage \cite{Fatimy2010, Rabbaa2011}. In this paper, a nanometer-scale imaging technique is proposed in which subwavelength plasmons generated by a current in the transistor channel, that can be tuned by controlling gate voltage, form the illumination pattern required for SIM. The configuration effectively creates a much larger observable spatial frequency region as compared to a freespace far-infrared (terahertz) wave. Due to the linear nature of the scheme, resolution of up to two orders of magnitude beyond the diffraction limit can be obtained with a weak field intensity, resulting in minimum damage to the sample. In contrast to the metal based SIM where the plasmonic wavelength is fixed by a determined structure \cite{Wei2010}, we can control the period of the plasmonic pattern by varying the gate voltage. Additionally, in comparison with graphene based SIM \cite{Zeng2014,Zeng2017}, our scheme uses surface current to excite the plasmons, which does not require any wavevector matching mechanism, such as gratings, and can be realized in experiment more easily.

A schematic diagram of the proposed system which is similar to a transistor, is shown in Fig. \ref{fig:struct} where a 2DEG, that acts as a transistor channel is formed at the interface of two semiconductor materials with slightly different band-gap energies. Plasmons are generated in the channel when the source and drain terminals are driven by a current source. Due to reflections from the conducting boundaries, a cavity in the channel region is created and the plasmons therefore, form a standing wave. The structure is backed by a gate terminal that spans the length $L$ of the channel and is spaced a distance $d$ below the 2DEG. The gate capacitively couples with the 2DEG, and by varying the voltage, the velocity as well as the concentration of electrons in the channel can be controlled. A barrier layer of thickness $h$ separates the sample from the 2DEG.

\begin{figure}[!hb]

\includegraphics[width=8cm]{fig1.eps}

\caption{(Color online) Illustration of the imaging scheme where the sample is excited by a standing plasmonic wave pattern generated in the 2DEG by a current-driven instability. S, D and G denote the source, drain and gate terminals of the transistor, respectively.}

\label{fig:struct}

\end{figure}

%

The compressed nature of the plasmons can be described by its dispersion relation. Here, we consider the sample stage which is terminated by a gate at the bottom and free-space at the top \cite{Kastner\_1988,Michalski2005}. The 2DEG is modeled as a shunt admittance related to a Drude-type surface conductivity \cite{Burke2000}, $\sigma\_{s} = N\_{s} e^{2} \tau/[m^{\ast}(1 -i \omega\_{p} \tau)]$, where $N\_{s}$ is the surface electron density in the channel, $e$ is the electron charge, $m^{\ast}$ is the effective electron mass in the heterostructure, $\tau$ is the scattering time of electrons, and $\omega\_{p}$ is the angular frequency. Through the gate voltage $V\_{g}$, the electron density $N\_{s}$ of the channel can be varied as $N\_{s} = N\_{0} \times (1 - V\_{g}/V\_{T} )$, where $N\_{0}=\varepsilon\_{2}\varepsilon\_{0}V\_{T}/(ed)$ is the zero-bias density, $V\_{T}$ is the gate threshold voltage of the transistor and $\varepsilon\_{2}$ is the relative permittivity of the substrate. The plasmonic dispersion relation can be written as $1-r\_{u}r\_{d}e^{-2k\_{p}h}=0$. Here, $r\_{u}$ accounts for the reflection from the barrier layer to vacuum while $r\_{d}$ represents the reflection from the barrier to the substrate with the 2DEG inbetween. Fig. 2 shows that the plasmonic wave number can be controlled by varying the gate voltage. The result is similar to the dispersion curve of gated 2DEG plasmons in which the gate terminal is located at the top \cite{Nakayama1974, Eguiluz1975}. Here, the permittivity of both semiconductor layers is approximated to the static value, i.e., $\varepsilon\_{1} \approx \varepsilon\_{2} = 9.5$; the plasmon frequency is $10~THz$; the mole-fraction of aluminum in AlGaN alloy is $0.1$ and the scattering time $\tau$ is $1.14~ps$ corresponding to a temperature of $77~K$ \cite{Muravjov2010}. As the temperature increases, $\tau$ reduces which leads to reduced mobility and introduces larger loss in the channel.

%

\begin{figure}

\includegraphics[width=7cm]{fig2.eps}

\caption{Plasmon wave dispersion diagram for a transistor structure supporting a 2DEG channel. Here, $V\_{T}=-9V$, $h=20nm$ and $d=120nm$.}

\end{figure}

%

As discussed before, the 2DEG channel essentially behaves as a cavity due to the resonance effects introduced by the two conducting boundaries, i.e., drain and source terminals. Therefore, the plasmonic wavenumber, $k\_{p}$ cannot be varied in a continuous fashion to cover all spatial frequencies. As an example, we consider the length of the heterostructure and the resulting 2DEG channel to be $2~\mu m$. Like any resonating structure, the plasmonic modes of the 2DEG channel are well-defined \cite{POPOV2007,Popov2008,Muravjov2010}. Fig. 3 shows simulations of plasmonic standing waves generated by a current-driven instability in the plasma channel using Comsol. In the simulations, the effective dielectric function of the 2DEG, which can be expressed as $\varepsilon(\omega\_{p}) = \varepsilon\_{s} -j \sigma\_{s} /(\omega\_{p} \Delta )$ \cite{Ando1982}, where $\Delta$ is the 2DEG thickness, is set to $2~nm$. In Fig. 3, it is shown that the standing wave pattern can be obtained and shifted by the surface current and an additional plane wave incident from above with an angle $\chi$ respect to the $z$ axis. The phase shifted field patterns, induced by the interference between the plasmons and the incident waves \cite{Zeng2017}, have the same period. Moreover, the period also can be tuned by varying the gate voltage.

%

\begin{figure}

\includegraphics[width=7cm]{fig3.eps}

\caption{Full-wave simulation results: Standing plasmonic wave patterns generated by a current-driven instability and an additional illumination at an angle $\chi$. The field intensities are normalized. The surface current intensity is $1~mA/m$ and the incident wave amplitude is $0.5~V/m$. The black solid and blue short dashed curves have incident angles $\chi=\pi/2~rad$ while the red dashed curve has an incident angle $\chi=0.245~rad$. The corresponding gate voltages of the black solid, red dashed and blue short dashed curves are $-5$, $-5$ and $0.4V$, respectively.}

\end{figure}

%

\begin{figure}

\includegraphics[width=7cm]{fig4.eps}

\caption{(Color online) Diagram of the spatial frequency. The blue and orange circles at the origin contribute to the conventional microscopy with imaging frequency $\omega\_{p}$ and $\omega\_{ac}$ respectively. The circles with the center at $\pm k\_{p 1/2}$ correspond to the linear response of the illumination pattern with period $2\pi/k\_{p 1/2}$. The period of illumination plasmons can be tuned by the gate voltage. (b) The energy structure of the sample molecule.}

\end{figure}

In linear SIM based on surface plasmons, the illumination pattern $I(\bm{r})$ can be assumed sinusoidal and expressed as $I(\bm{r}) = C + (e^{i\bm{k}\_{p} \cdot \bm{r}+i\phi}+e^{-i\bm{k}\_{p} \cdot \bm{r}-i\phi})/2$, where $C$ is a constant, attributed to the background, $\bm{k}$ is the spatial frequency wavevector, and $\phi$ is the pattern phase. An image $M(\bm{r})$ of a sample atom distribution $f(\bm{r})$ observed through a microscope can be expressed as,

%

\begin{equation}

M(\bm{r}) = [F(\bm{r}) \cdot I(\bm{r})] \otimes H(\bm{r}),

\end{equation}

%

where $H(\bm{r})$ is the point spread function (PSF) of the microscope, and $\cdot, \otimes$ denote multiplication and convolution operations in the spatial domain respectively. A spatial frequency domain representation of the image obtained by taking the Fourier transform is expressed as $\tilde{M}(\bm{k})=\tilde{F}(\bm{k})\tilde{H}(\bm{k})$, where $\sim$ indicates a spatial frequency domain term. Here, $\tilde{H}(\bf{k})$ is the optical transfer function (OTF) of the microscope, and $\tilde{F}(\bm{k})$ is the Fourier transform of $F(\bm{r}) \cdot I(\bm{r})$. A spatial frequency representation of the scheme is illustrated in Fig. 4(a). In this scheme, we assume that the numerical aperture of the objective lens is unity. In conventional fluorescence microscopy, the observable spatial frequency is limited to a circle as shown in Fig. 4(a) where the passband is bounded by $(k\_x^2 + k\_y^2)^{1/2} = 2 k\_{0} =2\omega\_{p}/c= \nu$. The plasmon frequency $\omega\_{p}$ falls in the terahertz region, therefore, a relatively small $k\_0$ implies that the image needs to be sampled a large number of times resulting in an imaging, which is slow. The process can be expedited by using an additional illumination such as a laser with frequency $\omega\_{e}$ in the visible region. As shown in Fig. 4(b), the molecules in the sample are first excited from the ground state, $|g\rangle$ to the energy level $|e\rangle$ by using a laser of frequency $\omega\_{e}$. The plasmonic field then excites the molecules to an additional level $|a\rangle$. By utilizing the spontaneous decay of the molecules from $|a\rangle$ to $|c\rangle$, we image the sample with photons of frequency $\omega\_{ac} = \omega\_{a}-\omega\_{c}$. A frequency-selective photonic crystal placed behind the objective lens can filter the photons of different frequencies \cite{Gustafsson2005}. As a consequence of the preceding discussion, the resulting passband in the spatial frequency is now bounded by $(k\_x^2 + k\_y^2)^{1/2} = 2 k\_{ac} = 2 \omega\_{ac}/c = \kappa$, which is the larger circle shown in Fig. 4(a). Since $\kappa$ is much larger than $\nu$, high resolution can be realized by imaging the sample only a few times.

A sinusoidal illumination pattern has three frequency components which, along with two shifted versions as shown in Fig. 4(a), generate an image from a linear combination of this frequency information. To reconstruct the sample, three different images need to be captured, each possessing a different phase term $\phi$. The process can be expressed as a system of linear equations:

%

\begin{equation}

\tilde{F}(\bm{k})=C \cdot \tilde{f}(\bm{k})+\tilde{f}(\bm{k}-\bm{k\_{p}})e^{i\phi}+\tilde{f}(\bm{k}+\bm{k\_{p}})e^{-i\phi}.

\end{equation}

%

The phase shifts $\phi$ are known beforehand. Frequency content of the sample up to $k\_{p}+\kappa$ can therefore be observed due the \emph{Moir$\acute{e}$} effect which transports the high frequency information into the observation region.

In order to obtain the three components of the spatial frequency as shown in above equation, we need to shift the plasmonic patterns. Our simulation results show that an additional incident plane wave can shift the pattern efficiently by varying its angle of incidence. In Fig. 3, the red and blue dashed curves are the shifted standing waves. To achieve two-dimensional enhancement in resolution, either the sample or the illumination above it must be rotated about the optical axis.

%

\begin{figure}

\includegraphics[width=7cm]{fig5.eps}

\caption{The Fourier components of the red dashed line in Fig. 3. $\Lambda=2\pi/S$ and $S=1\mu m$ is the size of the sample along $x$.}

\end{figure}

%

\begin{figure\*}

\centering

\includegraphics[width=15cm]{fig6.eps}

\caption{(a) Sample distribution. Simulation of the reconstructed sample image at different parameters: (b) $Re(k\_{p}^{max})=160k\_{0}$ (c) $Re(k\_{p}^{max})=360k\_{0}$.}

\end{figure\*}

%

Generation of plasmons with separated wave numbers is accomplished due to the resonator effect in the transistor channel. The mode separation can be approximated as $\Delta\_{k} =2 \pi\times10^{6} rad/s$. Therefore, just by tuning the plasmonic wavelengths, full coverage of all the spatial frequencies cannot be accomplished. However, as discussed earlier, through the larger circle shown in Fig. 4(a), all the spatial frequencies up to the plasmonic wavenumber can be recovered. For a laser, $\omega\_{e}$ having a wavelength of 600nm, the circle radius is $\kappa =4\pi/6\times 10^{7}rad/s$. We note that since $\Delta\_{k} < \kappa$, all the spatial content up to $\kappa$ can therefore be obtained.

The irregularities of the field pattern illustrated in Fig. 3 are contributed by the boundaries of finite conductivity and loss in the 2DEG. Therefore, the image cannot be simply resolved into just three spatial components as expressed in Eq. (2). However, as proved in Ref. 5, if the sample size or even the transistor channel length is much smaller than the plasmonic wave propagation length, i.e., $1/Im(k\_{p})$, the loss resulting from irregularities can be neglected. With ever improving nanofabrication processing techniques, the loss along the 2DEG channel can be very small, which subsequently means a large $Re(k\_{p})/Im(k\_{p})$ ratio. Additionally, in order to improve the accuracy of the spatial frequency components, we can borrow the mathematical process in Ref. 6 where the field intensity is Fourier expanded and its distribution is limited by the plasmon wave propagation length, as shown in Fig. 5 in which the sample size is chosen as $1~\mu m$. The figure shows that only a few Fourier components are large enough to overcome the noise. As a consequence, we still can solve for the spatial frequencies just by imaging a few times.

With all the Fourier information now known, an image of the sample can be reconstructed. We consider a sample with an atom distribution as shown in Fig. 6(a). The minimum separation between the atoms is $38~nm$ and maximum is $137~nm$. The 2DEG plasmons have a frequency of $\text{10~THz}$ and the numerical aperture (NA) is assumed to be $1$. The reconstruction involves contributions from the spatial frequency content up to a circular region of radius $k\_{p} + \kappa$, where $k\_{p}$ can be varied by gate voltage. In Figs. 6(b) and 6(c), reconstructed images with maximum plasmonic wavenumber $160$ and $360$ times larger than that of freespace are shown, corresponding to resolutions of 150 nm and 75 nm, respectively. Fig. 6(b) shows that the particles that are separated by a distance less than the resolution cannot be resolved and appear as a contiguous blurry streak, whereas they are distinguishable in Fig. 6(c).

Due to the large circle radius $\kappa$ shown in Fig. 4(a), a total of only 30 to 50 images are required for the final image reconstruction, which in terms of imaging speed is very fast and is comparable to nonlinear SIM \cite{Gustafsson2005}. However, unlike non-linear SIM, the method described here requires only illumination intensity.

In this paper we have proposed a super-resolution microscopy scheme based on subwavelength surface electromagnetic plasmons found in a semiconductor heterostructure. This method is useful in particular for light-sensitive samples as it requires weak field intensity for illumination, which can be used in bioimaging. Compared to metal based SIM where the plasmonic wavelengths are fixed, the period of the plasmonic pattern can be controlled by varying the gate voltage. Moreover, compared to graphene based SIM \cite{Zeng2014}, our scheme uses surface current to excite the plasmons, no complex wavevector matching structure is needed and it can be realized in experiment.

\begin{acknowledgments}

This research is supported by National Priorities Research Program (NPRP) Grant No. 8-352-1-074 from the Qatar National Research Fund (QNRF) .

\end{acknowledgments}